

L'évaluation des apprentissages en contexte scolaire Développements, enjeux et controverses

Lucie Mottier Lopez et Dany Laveault

Volume 31, numéro 3, 2008

URI : <https://id.erudit.org/iderudit/1024962ar>

DOI : <https://doi.org/10.7202/1024962ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (imprimé)

2368-2000 (numérique)

[Découvrir la revue](#)

Citer cet article

Lopez, L. M. & Laveault, D. (2008). L'évaluation des apprentissages en contexte scolaire : développements, enjeux et controverses. *Mesure et évaluation en éducation*, 31(3), 5–34. <https://doi.org/10.7202/1024962ar>

Résumé de l'article

La recherche sur l'évaluation en contexte scolaire a considérablement augmenté et diversifié ses objets d'étude depuis les travaux docimologiques sur les examens et les travaux sur l'évaluation formative de Bloom et ses collègues dans les années 1970. L'article expose les principaux élargissements du champ de recherche et discute les réaménagements et redéfinitions des concepts et notions tels qu'ils ont été notamment présentés dans la revue *Mesure et évaluation en éducation*. Ainsi sont discutés le rôle important de la définition des objectifs d'apprentissage, les élargissements du modèle initial de l'évaluation formative, les différents buts de l'évaluation sommative, l'enjeu majeur de l'évaluation de compétences. Malgré des avancées significatives réalisées dans le domaine, l'article montre que le besoin de recherches est encore plus pressant que jamais alors que l'évaluation scolaire doit s'inscrire dans un contexte de réforme, d'attentes et d'exigences de plus en plus nombreuses.

L'évaluation des apprentissages en contexte scolaire : développements, enjeux et controverses

Lucie Mottier Lopez

Université de Genève

Dany Laveault

Université d'Ottawa

MOTS CLÉS: Approche par compétences, autoévaluation, évaluation certificative, évaluation formative, objectifs d'apprentissage

La recherche sur l'évaluation en contexte scolaire a considérablement augmenté et diversifié ses objets d'étude depuis les travaux docimologiques sur les examens et les travaux sur l'évaluation formative de Bloom et ses collègues dans les années 1970. L'article expose les principaux élargissements du champ de recherche et discute les réaménagements et redéfinitions des concepts et notions tels qu'ils ont été notamment présentés dans la revue Mesure et évaluation en éducation. Ainsi sont discutés le rôle important de la définition des objectifs d'apprentissage, les élargissements du modèle initial de l'évaluation formative, les différents buts de l'évaluation sommative, l'enjeu majeur de l'évaluation de compétences. Malgré des avancées significatives réalisées dans le domaine, l'article montre que le besoin de recherches est encore plus pressant que jamais alors que l'évaluation scolaire doit s'inscrire dans un contexte de réforme, d'attentes et d'exigences de plus en plus nombreuses.

KEY WORDS: Competency-based instruction, self-assessment, certification, formative assessment, learning objectives

Since the 70s seminal work of Bloom and his colleagues on formative assessment, classroom assessment research has considerably enlarged and diversified its investigations. This paper illustrates the main directions of the field's expansion and addresses some aspects of the reorganization of the notions and concepts as they occurred in the journal Mesure et évaluation en éducation. The paper focuses on the important role of learning objectives, on the growth of formative assessment from its original model, on the distinction to be made between different purposes of summative assessment and finally on the major issue around the assessment of competencies. Despite significant progress in the field, this paper concludes that classroom assessment research has never been so much needed as it must meet with an increased number of requests and requirements initiated by the recent school reforms.

PALAVRAS-CHAVE: Abordagem pelas competências, auto-avaliação, avaliação certificativa, avaliação formativa, objectivos de aprendizagem

*Na sequência dos trabalhos docimológicos sobre os exames e dos trabalhos sobre avaliação formativa realizados por Bloom e seus colegas durante os anos 70, a investigação em avaliação em contexto escolar aumentou e diversificou os seus objectos de estudo. Este artigo expõe os principais alargamentos do campo da investigação e discute a reestruturações e as redefinições dos conceitos e das noções tal como surgiram, nomeadamente, na revista *Mesure et évaluation en éducation*. Neste sentido, propõe-se a discussão do papel importante da definição de objectivos de aprendizagem, dos alargamentos do modelo inicial de avaliação formativa, das diferentes finalidades da avaliação sumativa e, finalmente, do desafio maior da avaliação de competências. Apesar de todos os progressos significativos realizados neste campo, o artigo mostra que a necessidade da investigação nunca foi tão urgente, na medida em que a avaliação escolar deve inscrever-se num contexto de reforma, de expectativas e de exigências cada vez mais numerosas.*

Notes des auteurs – Toute correspondance peut être adressée comme suit : Lucie Mottier Lopez, FPSE, Uni-Mail, 40, Bd Pont d'Arve, 1205 Genève, ou Dany Laveault, Faculté d'éducation, Université d'Ottawa, Pavillon Lamoureux, bureau 469, 145, rue Jean-Jacques Lussier, Ottawa, Ontario, K1N 6N5, ou par courriel aux adresses suivantes : [Lucie.Mottier@unige.ch] ou [dlaveaul@uottawa.ca].

Dès les années 1930-1940, les études docimologiques¹ ont mis en avant, des deux côtés de l'Atlantique, des dysfonctionnements dans les procédures de notation en situations d'examen. Cette perspective docimologique a représenté une première orientation scientifique étudiant l'évaluation scolaire dans sa fonction d'attestation des apprentissages des élèves par le moyen des examens. Elle n'avait pas pour projet d'examiner les pratiques évaluatives en contexte de classe, ni de différencier les apprentissages sur le plan individuel. C'est essentiellement avec les travaux de Bloom, Hastings et Madaus (1971) qu'un changement de regard sur l'évaluation scolaire s'opère: l'évaluation apparaît liée à des enjeux d'apprentissage et plus seulement de certification et de sélection sociale. On s'intéresse alors à «l'activité évaluative» et à son rôle capital dans la régulation des processus d'apprentissage et d'enseignement et également dans la régulation des systèmes de formation. Les objets d'étude portent notamment sur la nature des objets à apprendre et à évaluer, les démarches et les outils à développer, les postures des acteurs et leur implication, les processus et effets produits.

En augmentant et en diversifiant ses objets d'étude, la recherche sur l'évaluation en contexte scolaire a subi d'importants changements qui se sont traduits par un réaménagement et une redéfinition des concepts et notions. Cet article tente de rendre compte de ces transformations, ainsi que des enjeux et controverses qui ont traversé l'évaluation scolaire au cours des trente années qui se sont écoulées depuis la fondation de la revue *Mesure et évaluation en éducation* (ci-après la revue). Cette tâche est particulièrement problématique, considérant l'évolution des idées et de la terminologie sur une aussi longue période de temps et sur deux continents caractérisés par des environnements scolaires à la fois similaires et différents. Notre recension des écrits se fonde sur les articles publiés dans la revue, que nous avons mis en relation avec des synthèses déjà existantes et des écrits similaires publiés dans le domaine de l'évaluation scolaire². La première partie de l'article rend compte de travaux qui ont discuté le rôle important de la formulation des objectifs d'apprentissage. La deuxième partie porte sur les élargissements du modèle de l'évaluation formative par rapport aux propositions initiales de Bloom et al. (1971). La partie suivante a pour objet les notions d'évaluation sommative et certificative et la quatrième partie porte, quant à elle, sur l'évaluation des compétences. Un tel effort de synthèse, à la fois historique et théorique, ne saurait être complet sans relever les défis auxquels se trouve présentement confrontée la recherche en évaluation scolaire. En guise de conclusion, nous verrons que malgré des avancées significatives réalisées dans le domaine, le

besoin de recherches visant à valider les conceptions théoriques est encore plus pressant que jamais, alors que l'évaluation scolaire doit s'inscrire dans un contexte de réforme, d'attentes et d'exigences de plus en plus nombreuses.

Trente ans d'évaluation scolaire de 1978 à 2008 : le rôle fondamental des objectifs

Publiée pour la première fois en 1978, la revue est alors contemporaine d'une réforme importante des programmes scolaires, caractérisée par la définition de programmes d'études en termes d'*objectifs d'apprentissage* et par une pédagogie axée sur l'atteinte et la maîtrise de ces objectifs. Cette réforme fait notamment écho aux travaux de Bloom et des publications de langue anglaise dans le domaine. La période qui nous intéresse s'achève avec l'introduction d'une autre réforme caractérisée cette fois-ci par une redéfinition des objectifs des programmes d'études en termes de *compétences* (voir quatrième partie de l'article). L'approche par objectifs consacre le détachement de l'évaluation scolaire d'une logique de classement des individus pour la rapprocher d'une logique d'apprentissage. Même si les définitions d'objectifs sont nombreuses et diversifiées, le fait même de situer l'élève par rapport à ses apprentissages plutôt que par rapport à ses pairs constitue une révolution copernicienne caractérisée par une évaluation critériée plutôt que normative.

Tout au long de son histoire, mais principalement à ses débuts, on retrouve de nombreux articles de la revue sur les objectifs. Il est possible de les regrouper en trois grandes catégories :

1. élaboration de différents systèmes de classification d'objectifs d'apprentissage ;
2. impact de la communication des objectifs sur le rendement de l'élève ;
3. recherche d'une cohérence entre objectifs de formation et d'évaluation.

Élaboration de différents systèmes de classification d'objectifs d'apprentissage

Morissette (1986) publie dans la revue une traduction et une adaptation du modèle de Jenkins et Deno pour les objectifs pédagogiques. Ce modèle en quatre niveaux vise à présenter les objectifs en fonction de leur degré de spécificité et des rôles et responsabilités des différents intervenants de l'éducation à chaque niveau. Il permet de décrire une certaine communication

ou transmission des attentes entre les différents intervenants, et ce, jusqu'à l'élève lui-même. Ce n'est donc pas uniquement un modèle de classification des objectifs, mais aussi un cadre conceptuel pour la validation du contenu des épreuves scolaires.

Le grand nombre et la diversité des classifications d'objectifs ont conduit quelques auteurs à proposer de nouvelles synthèses (par exemple Tochon, 1989, dans la revue). D'autres efforts d'intégration se sont poursuivis du côté anglo-saxon, dont le plus ambitieux est sans doute le travail d'Anderson et Krathwohl (2001) consistant à réactualiser la taxonomie de Bloom en y introduisant les dimensions de la métacognition et de la créativité. Les auteurs ont également pour ambition de démontrer les interactions entre enseignement, apprentissage et évaluation. Cette recherche de cohérence et d'articulation des trois domaines est tout à fait caractéristique des évolutions les plus récentes de l'évaluation en éducation.

Impact de la communication des objectifs sur le rendement de l'élève

Les premières recherches de Mager et McCann (1961, cités par Morissette, 1986) rapportaient déjà que la transmission en détail des objectifs pédagogiques réduisait de façon significative le temps nécessaire à la formation des élèves. L'étude de Tourneur (1989) dans la revue confirme ces résultats et les précise au moyen d'un dispositif expérimental rigoureux. Les élèves qui ont disposé de la liste des objectifs opérationnalisés réussissent mieux que ceux qui n'ont pas reçu une telle liste. Tourneur démontre également que l'administration d'un prétest exerce un effet de sensibilisation sur la réussite d'égale efficacité à celui de la transmission des objectifs.

Les résultats de Tourneur (1989) sur l'effet du prétest sont confirmés indirectement par les résultats de Laveault et Fournier (1990) et de Fournier et Laveault (1994), qui démontrent que la capacité des élèves à anticiper les questions qui leur sont posées est un bon prédicteur de la réussite à un examen. S'inspirant du modèle théorique de la métacognition, les auteurs en concluent qu'une meilleure communication des objectifs améliore les processus de surveillance métacognitive et sont susceptibles d'accroître l'habileté à se préparer à un examen.

Le modèle métacognitif va rapidement se révéler insuffisant. Déjà Delorme (1981) dans la revue soulevait la question de l'importance de l'affectivité dans une pédagogie par objectifs et constatait que «l'annonce

d'objectifs à des élèves pouvait améliorer les conditions affectives de travail» (p. 79). Il s'interroge sur l'après-pédagogie par objectifs : «Et si c'était au niveau du désir et de la liberté d'apprendre qu'allaient se jouer ou se jouent les prochains et principaux conflits éducatifs?» (p. 79). Laveault (1999) précise le rôle joué par la perception du but à atteindre dans l'autorégulation du comportement. La perception d'efficacité personnelle, notamment, interagit avec les différents systèmes de perception de soi et intervient dans la prise en charge de l'action personnelle. Il ne suffit donc pas que l'élève perçoive les objectifs pour s'engager cognitivement dans une tâche, mais il faut aussi qu'il évalue positivement ses chances de réussite.

Recherche d'une cohérence entre objectifs de formation et d'évaluation

Selon Cardinet (1984), les objectifs sont d'abord l'expression d'un jugement de valeur : «[...] un type de résultats n'est pas plus "vrai" qu'un autre, mais simplement plus "utile"» (p. 17). L'auteur considère sept types d'objectifs selon leur fonction utilitaire et selon leur capacité à tenir compte des besoins de formation et des formés. Son découpage des objectifs se fait selon deux dimensions :

- l'objet: il peut s'agir de contenus (connaissances, notions d'un programme), de démarches de pensée (aptitudes ou capacités) et de savoirs en situation «qui se situent dans le domaine de l'action, sur le plan du réel» (p. 22);
- le point de vue: il peut être statique (un produit, des apprentissages terminés, les connaissances acquises, des qualifications-clés) ou encore dynamique (le processus d'apprentissage lui-même).

Cardinet (1984) constate que la catégorie d'objectifs la plus répandue consiste à combiner les contenus et les qualifications-clés. Par contre, la communication de ces types d'objectif représente de plus grands défis que ceux qui sont définis en termes de fonctions : «Montrer aux élèves les situations auxquelles ils doivent pouvoir faire face est plus mobilisateur pour eux qu'une liste de contenus ou de savoir-faire généraux» (p. 32). D'une certaine manière, les objectifs définis en termes de fonctions anticipent sur l'approche par compétence avec toutes les difficultés nouvelles qu'elle soulève. En effet, l'évaluation d'objectifs en termes de fonctions, tout comme les compétences, suppose que soit établie «une progression de situations» (Cardinet, 1984,

p. 33) ou «parcours de développement» (Tardif, 2006). Tout en admettant qu'il soit possible d'établir une telle progression, Cardinet reconnaît qu'il s'agit de la principale difficulté d'une approche fonctionnelle.

Vers une conception élargie de l'évaluation formative

Les propositions de la pédagogie de la maîtrise développées par Bloom et ses collègues (Bloom, 1968, 1976; Bloom et al., 1971) renversent la fonction de l'évaluation scolaire: celle-ci doit d'abord être au service des apprentissages des élèves (et de leur réussite scolaire) avant même de les attester. Malgré les réserves qui seront ensuite formulées à l'encontre des propositions de Bloom, il faut reconnaître à sa juste valeur le changement paradigmatique qu'elles opèrent: Bloom conceptualise explicitement l'évaluation au regard d'une théorie d'apprentissage, le modèle néo-behavioriste, il développe un modèle qui s'adresse aux enseignants et qui vise une transformation de leurs pratiques pédagogiques. Ainsi, les pratiques évaluatives dans le contexte de la classe deviennent objets d'investigations scientifiques, de développements théoriques, de propositions pratiques. Par exemple, les recherches de Bloom (1984) tendent à démontrer qu'un enseignement collectif peut être aussi efficace qu'un enseignement par tutorat. Ces travaux sont notamment repris dans la revue par Huberman, Juge et Hari (1985) qui en confirment l'essentiel tout en nuancant davantage leur portée.

Un enjeu important des travaux scientifiques a été d'élargir dans plusieurs directions le modèle initial de l'évaluation formative de Bloom. Nous discutons ci-dessous les principales orientations de cet élargissement en nous appuyant notamment sur la recension effectuée par Allal et Mottier Lopez (2005), qui ont examiné systématiquement les articles publiés dans la revue entre 1978 et 2002 sur cette question. Pour l'article, nous avons complété avec les numéros suivants jusqu'en 2008. Nous soulignons ci-dessous l'accent mis sur l'instrumentation de l'évaluation formative, sur la recherche de cadres théoriques de plus en plus élaborés et sur la production de savoirs spécialisés qui en a découlé. Cela nous amène à discuter le rôle capital de l'implication de l'élève dans les processus évaluatifs et à souligner l'importance des études des pratiques évaluatives en contexte.

Centration sur l'instrumentation de l'évaluation formative

Dans un premier temps, les chercheurs francophones se sont préoccupés de l'*instrumentation* de l'évaluation formative en prenant modèle sur le *Handbook on formative and summative evaluation of student learning* de Bloom et al. (1971). Des recueils d'instruments ont été publiés dans différents domaines disciplinaires. Dans la revue, plusieurs articles ont contribué au développement de cette instrumentation, par exemple pour définir des lignes directrices pour construire des épreuves critériées (Racine, 1982) ou encore pour proposer des épreuves assistées par l'informatique à des fins d'évaluation diagnostique des erreurs (Dassa, 1988; De Campos, 1990; Séguin, 1984). Une des critiques majeures formulée relativement à ces travaux est qu'ils nourrissent une forme «d'illusion instrumentale» (e.g., Bain, 1988, dans la revue) et risquent de faire l'économie d'une réflexion théorique sur les processus d'enseignement/apprentissage rattachés à l'évaluation formative. Notons que cette orientation sur l'instrumentation de l'évaluation est toujours restée très présente, tout en tentant de considérer les critiques formulées. Les travaux sur le portfolio sont de ce point de vue représentatifs, ancrés dans une réflexion instrumentale et théorique sur les conceptions d'apprentissage sous-jacentes à cet outil et à ses caractéristiques (Mottier Lopez, 2006b).

À la recherche de cadres théoriques de l'évaluation formative

Après les questions d'instrumentation, un enjeu important a été de développer des cadres théoriques pour approfondir la conceptualisation de l'évaluation formative. Nous citons ci-dessous trois développements particuliers parmi d'autres possibles.

Apports des théories psychologiques de l'apprentissage

Le premier, dans le prolongement des propositions de Bloom, puise dans les *théories psychologiques* de l'apprentissage. Allal (1979) démontre les différences et les implications importantes de l'approche psychologique adoptée pour l'évaluation formative. L'auteure contraste la conception néo-béhavioriste de Bloom et l'approche constructiviste et sociocognitiviste qui amène à privilégier des modalités d'évaluation formative centrées sur les processus cognitifs d'apprentissage et de régulation et non pas seulement sur les performances des élèves et leur remédiation. D'autres auteurs ont poursuivi ce type de réflexion dans une approche constructiviste (e.g., Crahay, 1986), Thouin (1982, 1993) dans la revue, tenant plus spécialement compte de la logique des contenus disciplinaires.

Cette façon de conceptualiser l'évaluation formative est toujours actuelle, en interaction avec le développement des travaux en psychologie de l'apprentissage³. Ainsi, au regard des théories socio-culturelles inspirées de Vygotsky, Allal et Pelgrims Ducrey (2000) avancent l'idée que l'évaluation formative interactive fournit un étayage à l'apprentissage situé dans la zone de développement proximal de l'élève. Plus récemment encore, les cadres théoriques de la cognition et de l'apprentissage «situés» sont exploités par Allal (2007a) et Mottier Lopez (2006a, 2006b), amenant à concevoir l'évaluation comme une forme de participation à la communauté d'apprentissage (Brown & Campione, 1995) que représente la classe. Dans cette conception, la participation des élèves à une évaluation dite «authentique» devient source d'apprentissage significatif. Il ressort de ces travaux que les façons de conceptualiser l'évaluation formative se diversifient et se précisent en fonction de la conception d'apprentissage.

Apports des didactiques des disciplines

Le domaine de la *didactique des disciplines* offre un cadre théorique pour étudier l'évaluation en tant que composante du système didactique qui lie, dans une relation triadique, l'enseignant, l'apprenant et le savoir à enseigner/apprendre (Bain, 1988 ; Chevallard, 1986 ; Garcia-Debanco & Mas, 1987). Cette approche théorique cherche à montrer comment la structuration des contenus des disciplines scolaires détermine les objectifs et les moyens de l'évaluation formative. Bain et Schneuwly (1993) considèrent qu'il importe d'identifier les «modèles de référence» relativement aux objets de savoir à enseigner, susceptibles d'orienter l'évaluation formative. Laveault (1992) affirme l'importance de tolérer des formes «asymétriques» de formation où les élèves ne font pas tous les mêmes apprentissages en même temps. Cette perspective nécessite des pratiques d'évaluation qui se diversifient en fonction des besoins des élèves, des choix didactiques et des objectifs pédagogiques qui en découlent.

Un des apports de l'approche didactique est de souligner les processus de négociation didactique entre les acteurs à propos des niveaux d'exigence, des règles et des résultats de l'évaluation, comme montré par Chevallard (1986). Dans la revue, il est intéressant de constater que Cardinet et Laveault (1996) choisissent le «triangle didactique» comme élément de structuration pour effectuer leur recension des dix années des travaux européens et nord-américains sur l'évaluation en éducation⁴. Ils y ajoutent le pôle «société» afin d'intégrer les recherches sociologiques. Les auteurs soulignent le «triomphe

de l'approche cognitive» qui a marqué les travaux sur l'évaluation. Ils mettent l'accent sur la dimension *compréhensive* des travaux scientifiques. Figari (2006) commente: «[...] il s'agit moins de mesurer, de vérifier que d'analyser des processus pour comprendre le sens des activités du sujet» (p. 8).

Approches pluridisciplinaires

D'autres auteurs argumentent une *approche pluridisciplinaire et systémique* pour étudier l'évaluation formative. Dans la revue, les articles de Perrenoud (1991, 1993) placent l'évaluation formative à l'intersection de plusieurs perspectives théoriques. Pour l'auteur, il est nécessaire d'articuler les orientations cognitives, communicatives et didactiques de l'évaluation formative dans un cadre plus général de la régulation. Ce cadre intègre les régulations découlant de l'évaluation formative mais il les dépasse pour englober l'ensemble des régulations en situation d'enseignement et d'apprentissage (Perrenoud, 1998). Pour leur part, Bonniol et Vial (1997) exploitent différentes implications des théories cybernétiques, systémiques et de la complexité pour penser l'évaluation formative. Ils montrent les implications théoriques et parfois pratiques des différentes orientations des modèles. Vial (2006), dans la revue, démontre les relations entre formation des enseignants et évaluation.

Production de savoirs spécialisés sur l'évaluation formative et la régulation

C'est Cardinet (1977) qui a introduit la notion de régulation, inspiré de l'analyse des systèmes cybernétiques, pour remplacer celle de «remédiation» proposée par Bloom. Le but était de ne pas réduire les gestes résultant de l'interprétation des données recueillies à des mesures correctives (remédiation), mais de les élargir à toute forme d'adaptation de l'enseignement à des fins de régulation des apprentissages. Sans viser l'exhaustivité, nous citons ci-dessous trois développements de la régulation en tant que cadre théorique devenu plus général.

Allal (1979, 1988, 1993) a très vite introduit des distinctions conceptuelles afin d'appréhender la temporalité de la régulation formative en contexte de classe (régulation *on line*, régulation différée), ses différentes modalités (interactive, rétroactive, proactive), ses niveaux d'organisation en situation scolaire (structures des situations d'apprentissage, interactions sociales avec l'enseignant, entre apprenants, processus d'autorégulation, usage d'outils). Allal (2007b) délimite un cadre conceptuel qui reprend, entre autres éléments,

ses conceptualisations initialement produites dans le cadre de l'évaluation formative mais pour les élargir à un cadre théorique plus général sur la régulation des apprentissages en situation scolaire et en formation.

Dans des ancrages théoriques autres que la psychologie de l'apprentissage, Bonniol et Vial (1997) mettent l'accent sur une «évaluation-régulation» et sur l'idée d'une régulation systémique et dynamique qui est susceptible de créer de nouvelles configurations d'apprentissage et d'interprétation intersubjective. Autrement dit, la régulation ne relève pas seulement, dans une logique de régularisation, d'une action de réajustement par rapport à une norme pré-établie. L'évaluation devient alors un système ouvert et créateur dans lequel le référentiel peut se construire, se (re)négocier et se transformer en cours d'actions conjointes (Figari, 2001).

Laveault (2007) pose plus particulièrement la question de l'auto-régulation. Il identifie deux visées complémentaires de l'autorégulation en situation scolaire. La première, *l'autorégulation des apprentissages*, concerne l'aptitude des élèves à prendre en charge leurs processus cognitifs et motivationnels pour apprendre. La deuxième, *l'apprentissage de l'autorégulation*, se préoccupe du développement même de cette aptitude à exercer un plus grand et meilleur contrôle sur l'apprentissage. Ces deux visées sont caractéristiques de la tension constitutive du champ qui nous intéresse, à savoir une (auto)évaluation qui est à la fois un moyen et un objet d'apprentissage.

Implication active de l'élève dans l'évaluation

Les travaux qui ont contribué à élargir la conception de l'évaluation formative remettent en question l'idée, véhiculée par le modèle de la pédagogie de la maîtrise, que c'est l'enseignant qui est le seul responsable de l'évaluation. Les auteurs argumentent désormais l'importance d'une implication active de l'élève dans les processus évaluatifs. Dans un état de la question sur le sujet, Laveault et Miles (2008) dans la revue ont fait valoir que toute approche qui peut rendre les critères d'évaluation plus spécifiques, moins ambigus, donc plus univoques, favorise l'efficacité des pratiques d'autoévaluation. Ces approches peuvent prendre plusieurs formes, dont le recours à des échelles descriptives ou encore à l'illustration de niveaux de performance au moyen de copies types. Le modèle de l'évaluation «formatrice» développé par Nunziati (1990) et Vial (1995) correspond au niveau d'engagement maximal du formé en impliquant celui-ci directement dans la définition et le choix des critères de performance qui serviront à l'évaluer.

Genthon (1991) va dans le même sens en insistant sur la construction d'une compréhension partagée des outils d'évaluation avec les élèves et d'une signification de l'évaluation au travers de «procédures participatives». Cette appropriation des critères d'évaluation peut se faire selon une démarche progressive, les grilles d'évaluation descriptives et les copies types pouvant servir dans un premier temps à étayer la compréhension des critères avant d'impliquer le formé dans leur élaboration de façon indépendante.

Sans s'inscrire explicitement dans le modèle de l'évaluation formatrice, Allal (1999) définit trois modalités d'implication de l'élève dans l'évaluation : l'autoévaluation au sens strict, l'évaluation mutuelle entre pairs et la coévaluation stipulant une confrontation des évaluations réalisées par l'enseignant et par l'élève. Un angle important d'étude s'ouvre sur les processus métacognitifs dans les processus d'évaluation formative et d'autoévaluation (Allal, 1993 ; Noël, 2006 ; Scallon, 1996) en lien avec les théories contemporaines de la psychologie cognitive. Laveault (1999) élargit la conceptualisation de l'autoévaluation en ajoutant les régulations motivationnelles en plus des régulations cognitives et métacognitives. Laveault (2005) définit l'autoévaluation comme un ensemble de pratiques d'auto-régulation qui se différencient par le niveau de contrôle exercé par le formé dans chacune des trois composantes de toute boucle de régulation : le choix des objectifs et des niveaux à atteindre, la recherche et l'utilisation du *feedback* sur ses résultats ou sa performance et le choix, plus ou moins autonome, des moyens d'action pour se rapprocher de la cible à atteindre. Il est donc possible de décrire un continuum variant d'une situation où l'autoévaluation est entièrement sous le contrôle du formé à une situation où celui-ci n'exerce qu'un contrôle minimal ou nul sur chacune des composantes.

Si aujourd'hui le rôle capital de l'autoévaluation est largement admis, sa mise en pratique ne va pas sans difficulté, d'une part, pour que les élèves apprennent à s'autoévaluer, d'autre part, au regard de la complexité du processus autoévaluatif qui suppose une prise de distance critique et la maîtrise d'un métalangage (Abrecht, 1991 ; Laveault, 2007). Il convient aussi pour Allal (1999) d'être attentif à l'existence de pièges et de dilemmes liés à l'autoévaluation des élèves. Laveault (2007), se basant sur le modèle de Baumeister, Heatherton et Tice (1994), observe que les régulations peuvent comporter des effets dysfonctionnels : elles peuvent être insuffisantes (*under-regulation*), fautives (*misregulation*) ou encore excessives (*overregulation*). Il est donc plus que jamais important de comprendre comment une régulation

«

échoue» et quelles sont les limites à une régulation «réussie».

Étude des pratiques existantes dans leurs contextes

Les travaux francophones sur l'évaluation formative, contrairement aux approches anglo-saxonnes (voir à ce propos la revue de documentation de Black & Wiliam, 1998), se caractérisent par des développements essentiellement théoriques, voire prescriptifs liés aux enjeux praxéologiques de l'évaluation en situation scolaire. Pour Abrecht (1991), «les présentations de l'évaluation formative "pèchent" régulièrement par abstraction et extériorité» (p. 123) sans prendre suffisamment en compte la multiplicité et la diversité des contextes de mise en œuvre.

Bien évidemment, des études empiriques ont été réalisées. Dans la revue, citons la recherche d'Allal (1983) qui a examiné l'interaction entre l'instrumentation et l'intuition des enseignants dans les pratiques de l'évaluation formative; celle de Weiss (1984) montrant l'incompatibilité entre certains instruments d'évaluation formative et les pratiques quotidiennes d'évaluation des enseignants; celle de Perrenoud (1991) qui s'est intéressée aux aspects systémiques de l'évaluation pouvant favoriser ou au contraire freiner le développement des pratiques d'évaluation formative; celle de Dassa et De Cotret (1993) qui a étudié les façons d'articuler les pratiques en situation de classe et des tests diagnostiques informatisés. Plus récemment, Mottier Lopez (2007b) a étudié les évaluations et régulations interactives guidées par l'enseignant telles qu'elles sont situées dans des dynamiques contrastées de microculture de classe et leur impact sur l'activité d'apprentissage de l'élève. Mais d'une façon générale, comme le souligne Allal (2006), davantage d'études empiriques systématiques sont à réaliser, notamment pour démontrer les effets de régulation résultant de dispositifs d'évaluation contrastés.

Nouveaux enjeux pour l'évaluation formative

Pour assurer le développement de la recherche sur l'évaluation formative et l'évolution des pratiques, il faudra d'abord sortir l'évaluation formative de l'aporie qui oppose «évaluation formative» et «évaluation sommative». Comme le font remarquer De Ketele (2006) et Figari (2006) dans la revue, cette opposition est malheureuse et confond les finalités de l'évaluation avec la démarche employée. L'évaluation formative décrit une finalité (supporter l'apprentissage), alors que l'évaluation sommative décrit une démarche (comment établir un bilan). Pour De Ketele (2006), il est préférable de différencier les formes d'évaluation selon leurs fonctions principales: «préparer une nouvelle action; améliorer une action en cours; certifier au terme d'une action» (p. 102)⁵. L'opposition entre sommatif et formatif a donné lieu

à un «réductionnisme historique [qui] a conduit à restreindre la démarche sommative à la fonction certificative, la démarche descriptive à la fonction régulatrice ou formative, la démarche herméneutique ou intuitive à la fonction d'orientation» (De Ketele, 2006, p. 103). 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. C'est ce qui ressort des travaux de Doyon et Juneau (1991), d'Allal, Wegmuller, Bonaiti-Dugerdil et Cochet Kaeser (1998) ou encore de Stiggins, Arter, Chappuis et Chappuis (2004). Le principal obstacle au développement de l'évaluation formative, cependant, ne vient pas seulement de cette confusion entre finalité et démarche, mais surtout de la carence de recherches empirique sur les déterminants et facteurs de son efficacité en salle de classe (Allal & Mottier Lopez, 2005). Il est difficile, dans ce contexte, de démontrer la validité des méthodes employées pour soutenir l'apprentissage là où les bonnes intentions ne suffisent pas. L'évaluation n'est pas formative si elle ne donne pas lieu à de nouveaux succès dans les apprentissages. Elle pose aussi avec acuité la question de la motivation des élèves.

Givvin, Stipek, Salmon et MacGyvers (2001) ont démontré que les enseignants perçoivent, bien souvent à tort, que la motivation de leurs élèves ne change pas au cours de l'année, alors que celle-ci varie considérablement selon les matières et le moment de l'année. Cette observation est suffisamment généralisée pour qu'on en conclue qu'il est primordial d'élargir le champ des interventions de l'enseignant : les régulations doivent également porter sur les déterminants de la motivation d'accomplissement de l'élève. L'ajustement des tâches et des situations pour tenir compte des différences individuelles ne doit pas se faire uniquement en considérant les facteurs de nature cognitive ou métacognitive, mais aussi au regard des facteurs d'ordre sociocognitif, motivationnel, émotionnel et affectif. Lorsqu'un élève éprouve des difficultés qui lui paraissent insurmontables, il peut parfois être plus approprié pour l'enseignant d'approfondir les raisons de son état affectif que de fournir à l'élève un *feedback* sur ses difficultés cognitives. Dans de tels cas, la capacité à reconnaître les émotions et à gérer avec sensibilité la relation avec l'élève en difficulté, deux caractéristiques majeures de l'intelligence émotionnelle (Goleman, 1998), peuvent s'avérer plus utiles que toute régulation de type cognitif ou métacognitif.

De l'évaluation sommative à l'évaluation certificative

Dans les travaux des premières heures, l'évaluation formative a représenté un objet d'étude important. Quant à l'évaluation sommative, elle a souvent été étudiée dans ses aspects ou effets négatifs. Tourneur (1985) évoquait déjà dans la revue «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)

Dans la revue, Robillard et Louis (1985) parviennent indépendamment au même constat et sont surpris de constater que le guide d'évaluation du Québec « concerne uniquement l'évaluation formative » (p. 24). Tout comme Tourneur, ils soulignent l'importance de se soucier de l'articulation entre les évaluations sommative et formative.

Pour Tourneur, le concept d'évaluation sommative est « polysémique » et un effort de clarification s'imposait déjà en 1985. Tout comme l'évaluation formative, l'évaluation sommative a été victime du « réductionnisme historique » décrit par De Ketele (2006). Au contraire de l'évaluation formative, cependant, l'évaluation scolaire à des fins de certification héritait d'une longue histoire, pas toujours glorieuse, de mauvaises pratiques de correction, de notation et d'examens. À ceci s'ajoutaient les excès de mesures quantitatives se traduisant par une perte ou une dénaturation de l'information à cause de pratiques inadéquates de codification, de transformation des notes ou d'une mauvaise utilisation des techniques statistiques.

C'est à travers leur impact sur les évalués que Trahan, Carbonneau, Robillard et Boisvert (1983) dans la revue proposent d'étudier les interventions évaluatives. La notation, par exemple, peut être utilisée pour motiver les évalués à bon ou mauvais escient : « La pratique courante de compter à peu près tout ce que les élèves font habitue l'élève à une motivation plutôt extrinsèque » (p. 77). Leur analyse psychologique fait ressortir que ce n'est pas tant l'évaluation (la notation, l'examen, la correction)

qui a des conséquences positives ou négatives, que l'usage plus ou moins approprié qui en est fait. Par exemple, un système d'évaluation où l'élève choisirait le moment de son examen développerait l'autonomie de l'élève et un style d'attribution causale interne.

La confusion et le manque de coordination entre les différentes finalités de l'évaluation ont donné lieu à des paradoxes toujours actuels. Martineau (1983), dans la revue, remet en question la procédure qui consiste à établir la note de l'élève à partir d'une moyenne de résultats accumulés sur une période de temps plus ou moins longue :

Pourquoi tenir compte d'une note faible qui remonte au début de l'apprentissage, si entre temps l'élève a démontré qu'il avait atteint l'objectif (ou les objectifs)? Cette procédure est pénalisante pour l'élève et de plus n'est pas conforme à l'esprit de l'évaluation formative. (p. 52)

La contradiction relevée par Martineau (1983) existe toujours, alors que le calcul de la moyenne des résultats de l'élève sur une longue période de temps est réintroduit dans le bulletin scolaire de plusieurs juridictions scolaires de langue française.

De cette rétrospective de l'évaluation certificative, il est possible de faire ressortir un certain nombre d'enjeux sur les plans de la démarche méthodologique, de sa finalité et de sa coordination avec les évaluations externes.

Sur le plan de la démarche méthodologique

Pour certifier, il faut nécessairement faire des bilans ; l'établissement de ceux-ci, qu'ils soient quantitatifs ou qualitatifs, pose toujours des défis considérables. Proposant de nouvelles bases pour la certification, Cardinet (1987) dans la revue fait valoir un certain nombre de solutions dont les examens sur mesure («*testing* adaptatif») pour tenir compte des différences individuelles dans la formation, le recours à des profils qualitatifs et à un portefeuille de compétences pour contrer la perte d'information liée à la codification. Pour la reconnaissance des acquis, Cardinet fait également valoir l'importance de référentiels détaillés permettant de préciser le niveau de compétence exigé.

Sur le plan de la finalité

L'évaluation certificative n'a pas à se dérouler à sens unique, de l'évaluateur à l'évalué. Doyon et Juneau (1991) proposent une autoévaluation en plusieurs phases qui se termine par la sélection d'objectifs personnels prioritaires pour chaque élève et d'objectifs collectifs pour l'ensemble de la

classe. L'élève est également impliqué dans le choix des différents outils de travail dont il aura besoin pour entreprendre la séquence suivante. Cette approche est similaire à l'approche de type entretien (*conference*) avec et à propos de l'élève que Stiggins et al. (2004) recommandent en plus du portfolio. Sur le plan conceptuel, il est intéressant de noter que ces auteurs considèrent le portfolio non pas comme un outil d'évaluation, mais comme un outil de communication des résultats qui regroupe et organise plusieurs outils d'évaluation tels que le journal de bord, des grilles d'appréciation, des échelles descriptives d'autoévaluation associés à des productions de l'élève également insérées dans le portfolio. Il offre une structure qui facilite la création de liens significatifs entre les différents éléments, intéressants pour un étayage à l'évaluation certificative. Dans la revue, Mottier Lopez (2006b) considère qu'un élément crucial est la construction interactive avec les élèves des pratiques et des significations associées à cet outil complexe. Il convient aussi de ne pas mésestimer les relations à construire avec les élèves entre l'usage qu'ils font de leur portfolio et les activités qu'il est sensé médiatiser.

Sur le plan de la coordination avec les évaluations externes

Même en présence de référentiels détaillés, *a fortiori* lorsque ceux-ci n'existent pas, la certification des acquis ou des résultats d'une formation donne toujours lieu à une certaine interprétation des évaluateurs et des formés. Dans la revue, Weiss (1994) reconnaît qu'une autre façon d'évaluer consiste également à :

[...] se doter de références provisoires et partielles [telles que] l'ensemble des résultats des classes d'une région [...]. Ces références mettent ainsi les enseignants en mesure de situer leurs appréciations et leurs degrés d'exigence ou de sévérité par rapport à des normes qui, à défaut d'être une liste d'objectifs critériés entendus au sens strict, n'en est pas moins une certaine expression. (pp. 69-70)

Au cours des 10 dernières années, plusieurs pays ou juridictions scolaires ont développé des épreuves externes, tant pour assurer une certaine régulation du système éducatif que pour permettre aux enseignants et aux établissements de mieux ajuster leurs interventions en fonction des exigences du curriculum. Un enjeu actuel est de renforcer la complémentarité entre les pratiques évaluatives des enseignants et les évaluations externes et d'améliorer chacune des formes d'évaluation en respectant leurs logiques propres (Allal, à paraître ; Laveault, à paraître a).

L'approche par compétences : évaluer un tout

L'approche par compétences est l'une des avancées les plus récentes dans le domaine de la définition d'objectifs de programme. La plupart des nouveaux curricula issus de la réforme des programmes d'études des pays francophones s'en inspirent à des degrés divers. Si la définition de compétence fait de plus en plus consensus, son opérationnalisation soulève des défis :

« [...] il est difficile de définir la notion de compétence indépendamment de son cadre d'apprentissage, et des situations prévues pour son évaluation » (IGEN, 2007, p. 12). L'opérationnalisation des compétences se traduit donc par des exigences nouvelles que nous aborderons dans les trois sections suivantes.

La planification de l'évaluation

La planification de l'évaluation des compétences relève d'une nouvelle logique (Jonnaert & Masciotra, 2007). Puisque les compétences décrivent des attentes d'apprentissage en termes de capacités à traiter une classe de situations, les référentiels de compétences doivent aller au-delà des programmes par objectifs et fournir aux enseignants des exemples concrets de situations et fournir aussi des prototypes d'activité. La planification pédagogique se fait « à rebours » (Ministère de l'Éducation de l'Ontario, 2006, 2008). Elle part des situations qui nécessitent la compétence pour élaborer et choisir toute une série de tâches qui permettront à l'élève d'effectuer les apprentissages prévus à sa formation. Selon la nature de la tâche d'apprentissage qui aura été travaillée et celle qui aura été choisie pour l'évaluer, l'élève peut soit restituer le savoir, comme dans des exercices d'application à l'identique, le réinvestir dans une situation familière ou encore dans une situation inédite (IGEN, 2007, p. 12).

L'opérationnalisation de cette « planification à rebours » soulève certains problèmes qui consistent d'abord à s'entendre sur le niveau de mobilisation. Scallon (2004) se demande jusqu'où on peut aller dans le choix de situations inédites pour susciter cette mobilisation. Crahay et Detheux-Jehin (2005), dans la revue, abordent la question du « savoir mobiliser » à travers l'étude de la résolution de problèmes dans différents contextes où la mobilisation des savoirs dépend de la surcharge de la mémoire de travail et du décodage du problème à résoudre.

La collecte des preuves de l'apprentissage

L'apprentissage de compétences étant davantage contextualisé, progressif et s'échelonnant sur de plus longues périodes de temps, la collecte de preuves d'apprentissage vise à suivre la progression de l'élève, d'abord pour la soutenir, mais aussi pour en rendre compte. Ces nouveaux besoins ont entraîné le développement de procédures originales de recueil d'information. Par exemple, Crahay et Detheux-Jehin (2005) observent la réussite des élèves à la fois dans des épreuves où l'énoncé « appelle explicitement l'exécution de certaines procédures » que dans des situations complexes où « il faut analyser l'énoncé et déduire la séquence de procédures qui conduit à la solution » (p. 63). Rey, Carette, Defrance et Kahn (2003) ainsi que Carette (2007) dans la revue proposent de procéder à la collecte de renseignements en trois phases. Dans la première phase, il s'agit de vérifier si l'élève sait « choisir et combiner correctement plusieurs compétences élémentaires pour traiter une situation nouvelle et complexe » (Carette, 2007, p. 60). Dans la seconde phase, il s'agit de vérifier si l'élève possède ces compétences et sait choisir celle qui convient dans une situation inédite. Enfin, la troisième phase cherche à confirmer si l'élève sait « exécuter une opération en réponse à un signal » (p. 61).

Les grilles d'observation, le portfolio et le journal de bord font également partie des méthodes de cueillette de données. Selon Dionne (2005) dans la revue, plusieurs de ces méthodes ne sont pas sans poser de nouvelles difficultés à l'évaluation certificative, notamment lorsque la situation nécessite un travail en équipe ou encore lorsqu'elle s'étend sur une longue période de temps.

L'évaluation de tâches complexes

Tant face à des finalités formatives que certificatives, l'évaluation de tâches complexes fait appel au travail d'équipe des enseignants et à l'exercice de leur jugement professionnel (Lafortune & Allal, 2008; Laveault, 2005; Laveault, à paraître b; Mottier Lopez & Allal, à paraître). Les élèves sont associés aux diverses étapes de l'évaluation pour en assurer la continuité mais aussi pour rendre compte de leur capacité à mobiliser leurs compétences de façon autonome et dans des situations nouvelles. L'évaluation de tâches complexes cherche donc à s'appuyer sur une plus grande variété d'instruments, de sources (élèves, enseignants, parents) et de stratégies. La fidélité des jugements est attestée tant au moyen de méthodes quantitatives (Scallon, 2004), qualitatives que mixtes (Laveault, à paraître b).

L'acquisition progressive des compétences et leur évaluation en continu font en sorte que le jugement de l'enseignant est appelé à s'exercer dans des conditions où les démarcations entre évaluation et apprentissage, entre évaluation formative et évaluation certificative, tendent à s'estomper. Pour Harlen (2005), il n'y a pas lieu de continuer à traiter séparément évaluation formative et évaluation certificative, car il existe une relation de réciprocité entre les deux que l'on aurait avantage à mieux exploiter. D'une part, l'évaluation formative, hautement critériée et contextualisée, possède un caractère volatil et peu fiable lorsque l'on suit de près les hésitations qui se produisent dans l'apprentissage continu de l'élève. D'autre part, le bilan des apprentissages offert par l'évaluation certificative, quoique plus stable, requiert une information riche comme celle recueillie dans le cadre de l'évaluation formative pour interpréter les résultats de l'élève et leur donner tout leur sens, particulièrement lorsque les décisions à prendre présentent des cas frontières et sont associées à des enjeux élevés.

Les méthodes de cueillettes de données utilisées par Rey et al. (2003) et par Carette (2007) illustrent la zone grise qui peut exister entre le formatif et le certificatif. D'une part, situer l'élève sur une échelle de progression qui s'étend de la capacité de l'élève à appliquer une procédure en réponse à une consigne jusqu'à la mobilisation de celle-ci dans une situation nouvelle, possède un intérêt pour la régulation de ses apprentissages. D'autre part, ces renseignements ont un intérêt certain quant à la progression scolaire de l'élève. L'évaluation certificative sert alors à déterminer si la position de l'élève sur une telle échelle de progression est suffisante pour reconnaître la compétence en question et pour la poursuite du cheminement régulier. De telles évaluations, tant à des fins formative que certificative et pronostique⁶, requièrent le développement d'un référentiel qui permette à l'enseignant de fonder son interprétation sur des valeurs communes et partagées, tant avec ses pairs qu'avec ses élèves. Il ne s'agit pas d'évaluer les apprentissages de l'élève au moyen d'un cumul d'indicateurs disparates, mais de les situer sur un «parcours de développement» (Tardif, 2006) tenant compte du curriculum et du programme d'études de l'élève.

Pour conclure

Les apports de la revue au champ de recherche sur l'évaluation scolaire

Nous avons choisi d'intégrer au fur et à mesure de l'article les publications qui ont contribué aux développements scientifiques et aux débats : leur discours est remarquablement contemporain et permet d'identifier des problématiques persistantes et en développement. Les contributions témoignent des principaux élargissements du champ et de la diversification et de la spécification des objets d'études. Par rapport aux premières années de la revue, il ressort que le nombre de publications se rapportant directement à l'évaluation en salle de classe est aujourd'hui nettement moins élevé, ce champ d'étude devant partager l'espace de la revue avec d'autres problématiques de l'évaluation dont se fait l'écho le numéro spécial actuel. Ceci contribue à placer l'évaluation scolaire dans une position toute relative. En effet, que peut-on dire de l'évaluation des apprentissages de l'élève lorsque l'évaluation des programmes d'étude ou du système éducatif fait apparaître des facteurs en partie à l'origine de ses réussites mais aussi de ses échecs ? Les articles de la revue témoignent du travail de clarification conceptuelle qu'il reste à faire pour offrir une base encore plus solide pour l'étude des pratiques évaluatives en contexte et pour la conception de dispositifs d'évaluation qui soient en phase avec les nouvelles conceptions d'apprentissage et théories didactiques. Nous concluons en citant ce qui représente, de notre point de vue, les principaux défis actuels.

Défis actuels

Nous avons montré que les travaux se sont attachés à définir des modèles de plus en plus sophistiqués de l'évaluation formative. Mais il convient d'être critique vis-à-vis de ceux-ci et de dépasser le discours des « bonnes intentions » (ou de passer de « l'utopie à la réalité » dirait Hadji, 1997) et de renforcer la validité écologique des études entreprises. Un défi pour les recherches futures est de mieux comprendre les dimensions *informelles* et *interactives* de l'évaluation au regard des normes, des pratiques, des significations coconstruites dans chaque microculture de classe (Mottier Lopez, 2007b, 2008) mais également entre les enseignants (Morrissette, 2008). Il s'agit également de mieux appréhender les aspects dysfonctionnels des régulations (Laveault, 2007) et d'oser questionner les possibles effets contre-productifs de l'évaluation sur l'apprentissage et la motivation des élèves.

L'évaluation tant formative que certificative doit faire place à la *professionnalité* des enseignants, s'inscrivant dans un rapport éthique et critique⁷ aux gestes professionnels dont fait partie le «savoir-évaluer» (Jorro, à paraître). Les études récentes sur le «jugement professionnel» (Lafortune & Allal, 2008 ; Laveault, 2005 ; Laveault, à paraître b) représentent une autre perspective pour comprendre comment les enseignants évaluent et prennent des décisions dans des situations pour lesquelles il n'existe pas de solution toute faite et qui souvent sont sources de dilemmes. Une culture d'évaluation⁸ plus collégiale et partagée, notamment par la mise en place d'équipes-cycles, de projets d'établissement, de fonctionnement des enseignants en communauté d'apprentissage, semble favorable à un changement de dynamique de l'évaluation scolaire, dans ses finalités multiples de soutien à l'apprentissage, de certification et d'orientation des élèves. Weiss (1994) dans la revue l'argumentait déjà en soulignant l'importance de l'attitude responsable et professionnelle des enseignants qui amènent à considérer l'évaluation, y compris certificative, comme un processus d'échanges et de communication favorable aux apprentissages et à la réussite des élèves.

Les travaux sur l'apprentissage et la cognition situés⁹ visent à examiner et à coordonner les nombreux aspects des relations entre évaluation, apprentissage et contextes (Allal, 2006, 2007a ; Mottier Lopez, 2006a). La thèse défendue est celle d'une relation de coconstitution entre les processus sociaux et culturels et les processus psychologiques individuels, chaque plan fournissant les conditions au développement et à la régulation de l'autre plan (Mottier Lopez, 2007a). Ce faisant, on s'intéresse à la «cognition en pratique» (*cognition in practice*, Lave, 1988), mettant l'accent sur la structuration dialectique entre l'activité cognitive et son contexte de réalisation, tant dans ses aspects sociaux, matériels que des significations construites. Un défi pour l'évaluation est de parvenir à prendre en compte cette relation entre apprentissages et contextes.

Afin de pouvoir remplir pleinement ses promesses, l'évaluation des attentes d'apprentissage formulées en termes de *compétences* devra poursuivre son développement tant sur le plan des stratégies que des instruments d'évaluation. Dans le rôle traditionnellement dévolu à l'évaluation formative, l'approche par compétences rend encore plus indispensable le recours à l'autoévaluation, non seulement en tant que processus d'autorégulation des apprentissages, mais aussi comme objet lui-même d'apprentissage (*assessment as learning*, Earl, 2003). Il importe de mieux comprendre aussi comment les

différentes modalités de reformulation des objectifs d'apprentissage (*deconstructing standards*, Stiggins et al., 2004) de même que le recours à des copies types permettent de clarifier les attentes et les niveaux d'exigence, de mobiliser les efforts de l'élève et d'en faire un partenaire actif du suivi de sa progression.

Dans son rôle d'évaluation certificative, la validité des décisions prises selon une approche par compétences dépend tant de la validité des résultats recueillis que de celle des échelles mises au point pour situer l'élève sur une trajectoire de développement. La construction de ces échelles fait appel au savoir professionnel des équipes d'enseignants et à leur capacité à comprendre et à mettre en pratique les grandes théories de l'apprentissage et de la didactique. La principale difficulté, lorsque les attentes sont définies en termes de « fonctions du savoir en situation », selon l'expression de Cardinet (1984, p. 21), ou « compétence » dans la terminologie actuelle, consiste à « établir un découpage des situations » et « à organiser une progression pédagogique à l'intérieur de chaque fonction » (p. 33). L'intégration de plusieurs types d'apprentissages (connaissances, compétences de base, compétences transversales et disciplinaires, etc.) s'avère particulièrement problématique lorsque plusieurs lignes de progression doivent s'articuler entre elles pour un développement harmonieux de différents types de savoirs.

La cohérence de l'évaluation certificative avec les fondements des programmes par compétences soulève également des questions lorsque les instruments utilisés mettent principalement l'accent sur une seule ou sur un nombre restreint de lignes de progression ou encore lorsque ceux-ci privilégient un seul ou un nombre réduit de conceptions de l'apprentissage. Il est tout à fait justifié d'exiger que soit démontrée la validité conceptuelle des résultats, surtout lorsqu'il y a des raisons de croire qu'ils ne permettent pas de faire les inférences souhaitées. Les principales critiques adressées à l'approche par compétences, ou encore aux instruments dans les enquêtes à grande échelle, vient de ce que, dans un cas comme dans l'autre, une conception simplifiée ou réduite des attentes d'apprentissage semble privilégiée au détriment d'autres. La diversification des pratiques et des objets d'apprentissage est nécessaire à l'exercice d'une pédagogie différenciée centrée sur les besoins de l'élève. Comme l'a démontré Cardinet (1984), opter pour un type d'objectifs, c'est faire un choix qui comporte des avantages et des inconvénients, qui répond à des besoins et correspond à un système de valeurs précis.

L'évaluation scolaire est à l'heure des choix et jamais l'enseignant n'aura disposé d'une telle palette de dispositifs d'évaluation. Pour faire des choix éclairés et judicieux, celui-ci est appelé à savoir mobiliser un ensemble grandissant de capacités et de ressources qui vont au-delà de l'application d'un savoir-faire technique. Au moment où nous attendons des élèves qu'ils démontrent leurs compétences dans des situations complexes et variées, la société et le système scolaire n'en attendent pas moins des enseignants en matière d'évaluation. Les défis pour les années à venir porteront tant sur l'évaluation des compétences et des connaissances des élèves que sur les connaissances et compétences professionnelles des enseignants à les évaluer.

NOTES

1. Les recherches docimologiques en tant que «science des examens» désignent un ensemble de travaux dans une approche métrique qui furent à l'origine de nombreuses interrogations sur les biais des systèmes de notation et des jugements évaluatifs.
2. Des parallèles seront établis entre notre recension des écrits majoritairement francophones et les travaux publiés en anglais dans le domaine. Nous ferons ressortir les points de convergence, tout en soulignant l'apport original des travaux francophones et plus spécifiquement ceux publiés dans la revue.
3. Les auteurs de ce courant mènent souvent conjointement des travaux sur les processus d'apprentissage et de régulation en situation scolaire.
4. Au sens large, c'est-à-dire pas seulement sur l'évaluation en contexte scolaire comme c'est le cas dans cet article.
5. Les deux dernières fonctions correspondent à la distinction établie par Stiggins, Arter, Chappuis et Chappuis (2004) entre «*assessment of learning*» et «*assessment for learning*».
6. Notre revue de questions n'aborde pas la fonction pronostique de l'évaluation qui est plus spécialement liée aux décisions de promotion/redoublement des élèves et aux décisions d'orientation, car très peu d'articles dans la revue en ont traité. On peut d'ailleurs le regretter car ce sont des situations d'évaluation qui se caractérisent par de forts enjeux, et qui mériteraient que des études examinent, notamment, leur relation avec les autres fonctions de l'évaluation.
7. Voir le numéro spécial de la revue sur «éthique et évaluation des apprentissages», volume 20(2) 1997.
8. Voir le numéro spécial de la revue sur «une culture de l'évaluation», volume 19(2) 1996.
9. Voir Mottier Lopez (2008) pour une revue de documentation.

RÉFÉRENCES

- Abrecht, R. (1991). *L'évaluation formative, une analyse critique*. Bruxelles : De Boeck.
- Allal, L. (1979). Stratégies d'évaluation formative : conceptions psycho-pédagogiques et modalités d'application. In L. Allal, J. Cardinet & P. Perrenoud (éd.), *L'évaluation formative dans un enseignement différencié* (pp. 153-183). Berne : Peter Lang.
- Allal, L. (1983). Évaluation formative : entre l'intuition et l'instrumentation. *Mesure et évaluation en éducation*, 6(5), 37-57.
- Allal, L. (1988). Vers un élargissement de la pédagogie de maîtrise : processus de régulation interactive, rétroactive et proactive. In M. Huberman (éd.), *Assurer la réussite des apprentissages scolaires ? Les propositions de la pédagogie de maîtrise* (pp. 86-126). Neuchâtel : Delachaux et Niestlé.
- Allal, L. (1993). Régulations métacognitives. In L. Allal, D. Bain & P. Perrenoud (éd.), *Évaluation formative et didactique du français* (pp. 81-98). Neuchâtel : Delachaux et Niestlé.
- Allal, L. (1999). Impliquer l'apprenant dans les processus d'évaluation : promesses et pièges de l'autoévaluation. In C. Depover & B. Noël (éd.), *L'évaluation des compétences et des processus cognitifs, modèles, pratiques et contextes* (pp. 35-56). Bruxelles : De Boeck.
- Allal, L. (2006). La fonction de régulation de l'évaluation : constructions théoriques et limites empiriques. In G. Figari & L. Mottier Lopez (éd.), *Recherche sur l'évaluation en éducation : problématiques, méthodologies et épistémologie* (pp. 223-230). Paris : L'Harmattan.
- Allal, L. (2007a). Évaluation dans le contexte de l'apprentissage situé. In M. Behrens (éd.), *La qualité en éducation* (pp. 39-56). Québec : Presses de l'Université du Québec.
- Allal, L. (2007b). Régulation des apprentissages : orientements conceptuels pour la recherche et la pratique en éducation. In L. Allal & L. Mottier Lopez (éd.), *Régulation des apprentissages en situation scolaire et en formation* (pp. 7-24). Bruxelles : De Boeck.
- Allal, L. (à paraître). Pratiques évaluatives des enseignants face aux méthodologies des recherches évaluatives portant sur le système scolaire. In L. Mottier Lopez & M. Crahay (éd.), *Évaluations en tension*. Bruxelles : De Boeck.
- 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).
- Allal, L., & Pelgrims Ducrey, G. (2000). Assessment of – or in – the zone of proximal development. *Learning and Instruction*, 10, 137-152.
- Allal, L., Wegmuller, E., Bonaiti-Dugerdil, S., & Cochet Kaeser, F. (1998). Le portfolio dans la dynamique de l'entretien tripartite. *Mesure et évaluation en éducation*, 20(3), 5-31.
- Anderson, L.W., & Krathwohl, D.R. (éd.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives*. New York : Longman.
- Bain, D. (1988). L'évaluation formative fait fausse route. *Mesure et évaluation en éducation*, 10(4), 23-32.

- Bain, D., & Schneuwly, B. (1993). Pour une évaluation formative intégrée dans la pédagogie du français: de la nécessité et de l'utilité de modèles de référence. In L. Allal, D. Bain & P. Perrenoud (éd.), *Évaluation formative et didactique du français* (pp. 51-79). Neuchâtel: Delachaux et Niestlé.
- Baumeister, R.F., Heatherton, T.F., & Tice, D.M. (1994). *Losing control: How and why people fail at self-regulation*. Toronto: Academic Press.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74.
- Bloom, B.S. (1968). Learning for mastery. *Evaluation Comment*, 1(2), 1-12.
- Bloom, B.S. (1976). *Human Characteristics and School Learning*. New York: McGraw-Hill.
- Bloom, B.S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6), 4-16.
- Bloom, B.S., Hasting, J.T., & Madaus, G.F. (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill.
- Bonniol, J.-J., & Vial, M. (1997). *Les modèles de l'évaluation*. Bruxelles: De Boeck.
- Brown, A.L., & Campione, J.C. (1995). Concevoir une communauté de jeunes élèves. Leçons théoriques et pratiques. *Revue française de pédagogie*, 11, 11-33.
- Cardinet, J. (1977). *Objectifs éducatifs et évaluation individualisée* (2^e éd., Rapport n° R77.05). Neuchâtel: Institut romand de recherches et de documentation pédagogiques.
- Cardinet, J. (1984). Le découpage des objectifs pédagogiques. *Mesure et évaluation en éducation*, 7(2), 5-48.
- Cardinet, J. (1987). De nouvelles bases pour la certification. *Mesure et évaluation en éducation*, 10(1), 5-25.
- Cardinet, J., & Laveault, D. (1996). Dix années de travaux européens et nord-américains sur l'évaluation: quelles lignes de force? *Mesure et évaluation en éducation*, 18(3), 1-25.
- Cardinet, J., & Laveault, D. (2001). L'activité éducative en éducation: évolutions des préoccupations des deux côtés de l'Atlantique. In G. Figari & M. Achouche (éd.), *L'activité évaluative réinterrogée. Regards scolaires et socioprofessionnels* (pp. 15-29). Bruxelles: De Boeck et Larcier.
- 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.
- Chevallard, Y. (1986). Vers une analyse didactique des faits d'évaluation. In J.-M. De Ketele (éd.), *L'évaluation: approche descriptive ou prescriptive?* (pp. 31-59). Bruxelles: De Boeck.
- Crahay, M. (1986). Évaluation formative et théorie constructiviste du développement. In J.-M. De Ketele (éd.), *L'évaluation: approche descriptive ou prescriptive?* (pp. 135-187). Bruxelles: De Boeck.
- Crahay, M., & Detheux-Jehin, M. (2005). L'évaluation des compétences, une entreprise impossible? *Mesure et évaluation en éducation*, 28(1), 57-78.
- Dassa, C. (1988). L'intégration du diagnostic pédagogique aux apprentissages scolaires: de la théorie à la pratique. *Mesure et évaluation en éducation*, 11(1), 7-26.

- Dassa, C., & De Cotret, S.R. (1993). Validation d'un système informatisé de diagnostic en mathématiques au secondaire : une approche centrée sur l'analyse didactique. *Mesure et évaluation en éducation*, 16(3-4), 5-26.
- De Campos, M. (1990). Outils diagnostiques et enseignement assisté par ordinateur. *Mesure et évaluation en éducation*, 13(3), 55-69.
- De Ketele, J.-M. (2006). La recherche en évaluation : propos synthétiques et prospectifs. *Mesure et évaluation en éducation*, 29(1), 99-118.
- Delorme, C. (1981). La pédagogie par objectifs ou quel changement de la pédagogie? *Mesure et évaluation en éducation*, 4(3), 55-60.
- Dionne, E. (2005). Enjeux en évaluation des apprentissages dans le cadre de programmes d'études formulés par compétences : l'exemple de la science et technologie. *Mesure et évaluation en éducation*, 28(2), 49-66.
- Doyon, C., & Juneau, R. (1991). *Faire participer l'élève à l'évaluation de ses apprentissages*. Laval, Québec : Beauchemin.
- Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks : Corwin Press Inc.
- Figari, G. (2001). Us et abus de la notion de référentiel. In G. Figari & M. Achouche (éd.), *L'activité évaluative réinterrogée, regards scolaires et socioprofessionnels* (pp. 310-314). Bruxelles : De Boeck.
- 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.
- Figari, G., & Achouche, M. (2001). *L'activité évaluative réinterrogée : regards scolaires et socioprofessionnels*. Bruxelles : De Boeck.
- Figari, G., & Mottier Lopez, L. (éd.). (2006). *Recherche sur l'évaluation en éducation : problématiques, méthodologies et épistémologie*. Paris : L'Harmattan.
- Fournier, C., & Laveault, D. (1994). Liens entre anticipation, autoévaluation et le résultat à un examen de rendement scolaire. *Revue des sciences de l'éducation*, 20(3), 427-442.
- Garcia-Debanc, C., & Mas, M. (1987). Évaluation des productions écrites des élèves. *Enjeux*, 11, 108-122.
- Genthon, M. (1991). Communiquer... quoi, pourquoi faire et comment, dans un processus d'évaluation. In J. Weiss (éd.), *L'évaluation : problème de communication* (pp. 175-188). Cousset : DelVal.
- Givvin, K.B., Stipek, D.U., Salmon, J.M., & MacGyvers, V.L. (2001). In the eyes of the beholder : Students' and teachers' judgments of students motivation. *Teaching & Teacher Education*, 17(3), 321-331.
- Goleman, D. (1998). *Working with emotional intelligence*. New York : Bantam Books.
- Hadji, C. (1997). *L'évaluation démystifiée*. Paris : ESF.
- Harlen, W. (2005). Teachers' summative practices and assessment for learning – tensions and synergies. *The Curriculum Journal*, 16(2), 207-223.
- Huberman, M., & Miles, M.B. (1991). *Analyse des données qualitatives*. Bruxelles : De Boeck.
- Huberman, M., Juge, P.-A., & Hari, P.-A. (1985). La pédagogie de la maîtrise : une évaluation instructive au niveau gymnasial. *Mesure et évaluation en éducation*, 8(3), 43-81.

- Inspection générale de l'éducation nationale (IGEN) (2007). *Les livrets de compétences : nouveaux outils pour l'évaluation des acquis*. Paris : IGEN, Rapport n° 2007-048.
- Jonnaert, P., & Masciotra, D. (2007). Sociocognitivism et logique de compétences pour les programmes d'études. In L. Lafortune, E. Moussadak & P. Jonnaert (éd.), *Observer les réformes en éducation* (pp. 53-76). Québec : Presses de l'Université du Québec.
- Jorro, A. (à paraître). L'évaluation comme savoir professionnel. In L. Mottier Lopez & M. Crahay (éd.), *Évaluations en tension*. Bruxelles : De Boeck.
- Lafortune, L., & Allal, L. (2008). *Jugement professionnel en évaluation*. Québec : Presses de l'Université du Québec.
- Lave, J. (1988). *Cognition in practice : Mind, mathematics and culture in everyday life*. Cambridge, MA : Cambridge University Press.
- Laveault, D. (éd.). (1992). *Les pratiques d'évaluation en éducation*. Montréal : Éditions de l'ADMEE.
- Laveault, D. (1999). Autoévaluation et régulation des apprentissages. In C. Depover & B. Noël (éd.), *L'évaluation des compétences et des processus cognitifs : modèles, pratiques et contextes* (pp. 57-79). Bruxelles : De Boeck.
- Laveault, D. (2005). Le jugement professionnel de l'enseignant : quel impact sur l'acte d'évaluer? *Mesure et évaluation en éducation*, 28(2), 93-114.
- Laveault, D. (2007). De la «régulation» au «réglage» : élaboration d'un modèle d'autoévaluation des apprentissages. In L. Allal & L. Mottier Lopez (éd.), *Régulation des apprentissages en situation scolaire et en formation* (pp. 207-234). Bruxelles : De Boeck.
- Laveault, D. (à paraître a). L'amélioration de l'efficacité du système éducatif : sur quels indicateurs s'appuyer? In X. Dumay & V. Dupriez (éd), *L'efficacité dans l'enseignement : promesses et zones d'ombre*. Bruxelles : De Boeck.
- Laveault, D. (à paraître b). Le jugement professionnel d'évaluation scolaire : enjeux, tensions et synergies nouvelles. *Revue suisse des sciences de l'éducation*, 3/2008.
- Laveault, D., & Fournier, C. (1990). L'évaluation par objectifs : une approche métacognitive. *Mesure et évaluation en éducation*, 13(1), 57-74.
- Laveault, D., & Miles, C. (2008). Utilité des échelles descriptives et différences individuelles dans l'autoévaluation de l'écrit. *Mesure et évaluation en éducation*, 31(1), 1-29.
- Martineau, G. (1983). Pour une évaluation actuelle. *Mesure et évaluation en éducation*, 6(2), 52-53.
- Ministère de l'Éducation de l'Ontario (2006). *Comment et pourquoi évaluer. Document d'appui en évaluation de la 1^{ère} à la 8^e année*. Toronto : auteur.
- Ministère de l'éducation de l'Ontario (2008). *Faire croître le succès*. Toronto : Auteur.
- Morissette, D. (1986). La mesure et l'évaluation du savoir être en classe : définition du problème et état des travaux sur une hypothèse de solution. *Mesure et évaluation en éducation*, 9(2-3), 5-38.
- Morissette, J. (2008). *Analyse des manières de faire l'évaluation formative des apprentissages d'un groupe d'enseignantes du primaire : une perspective inter-actionniste*. Thèse de doctorat non publiée, Université Laval, Québec.

- Mottier Lopez, L. (2006a). L'évaluation située des apprentissages: le rôle fondamental de la médiation sociale. In G. Figari & L. Mottier Lopez (éd.), *Recherche sur l'évaluation en éducation: problématiques, méthodologies et épistémologie* (pp. 82-89). Paris: L'Harmattan.
- Mottier Lopez, L. (2006b). Interroger la pratique du portfolio en situation scolaire dans une perspective «située» de l'apprentissage. *Mesure et évaluation en éducation*, 29(2), 1-21.
- Mottier Lopez, L. (2007a). Constitution interactive de la microculture de classe: pour quels effets de régulation sur les plans individuel et communautaire? In L. Allal & L. Mottier Lopez (éd.), *Régulation des apprentissages en situation scolaire et en formation* (pp. 149-169). Bruxelles: De Boeck.
- Mottier Lopez, L. (2007b). Régulations interactives situées dans des dynamiques de micro-culture de classe. *Mesure et évaluation en éducation*, 30(2), 23-47.
- Mottier Lopez, L. (2008). *Apprentissage situé: la microculture de classe en mathématiques*. Berne: Peter Lang.
- Mottier Lopez, L., & Allal, L. (à paraître). Le jugement professionnel en évaluation: un acte cognitif et une pratique sociale située. *Revue suisse des sciences de l'éducation*, 3/2008.
- Noël, B. (2006). La métacognition comme référence de l'autoévaluation. In G. Figari & L. Mottier Lopez (éd.), *Recherche sur l'évaluation en éducation: problématiques, méthodologies et épistémologie* (pp. 76-81). Paris: L'Harmattan.
- Nunziati, G. (1990). Pour construire un dispositif d'évaluation formatrice. *Cahiers pédagogiques*, 280, 48-64.
- Perrenoud, P. (1991). Pour une approche pragmatique de l'évaluation formative. *Mesure et évaluation en éducation*, 13(4), 49-81.
- Perrenoud, P. (1993). Touche pas à mon évaluation! Une approche systémique du changement. *Mesure et évaluation en éducation*, 16(1-2), 107-132.
- Perrenoud, P. (1998). From formative evaluation to a controlled regulation of learning processes: Towards a wider conceptual field. *Assessment in Education*, 5(1), 85-102.
- Racine, S.B. (1982). La validité et la fidélité dans la mesure critériée. *Mesure et évaluation en éducation*, 5(3), 92-110.
- Rey, B., Carette, V., Defrance, A., & Kahn, S. (2003). *Les compétences à l'école*. Bruxelles: De Boeck.
- Robillard, R., & Louis, R. (1985). Note de lecture. Critique du guide d'évaluation en classe au primaire. *Mesure et évaluation en éducation*, 7(4), 21-73.
- Scallon, G. (1996). Évaluation formative et psychologie cognitive: mouvances et tendances. In G. Grégoire (éd.), *Évaluer les apprentissages: les apports de la psychologie cognitive* (pp. 159-173). Bruxelles: De Boeck.
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Saint-Laurent, QC: ERPI.
- Séguin, S. P. (1984). L'utilisation des micro-ordinateurs pour l'évaluation des apprentissages: quelques perspectives d'avenir. *Mesure et évaluation en éducation*, 7(1), 53-64.
- 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.

- Tardif, J. (2006). *L'évaluation des compétences : documenter le parcours de développement*. Montréal : Chenelière-Éducation.
- Thouin, M. (1982). La définition des objectifs et l'évaluation dans le domaine affectif, quelques éléments de réflexion. *Mesure et évaluation en éducation*, 5(1), 31-34.
- Thouin, M. (1993). L'évaluation des apprentissages en mathématiques : une perspective constructiviste. *Mesure et évaluation en éducation*, 16(1-2), 47-64.
- Tochon, F. (1989). Vous avez dit trois niveaux d'objectifs? *Mesure et évaluation en éducation*, 11(4), 24-47.
- Tourneur, Y. (1985). La certification des compétences. *Mesure et évaluation en éducation*, 7(4), 5-20.
- Tourneur, Y. (1989). Les effets sur l'apprentissage de la communication des objectifs. *Mesure et évaluation en éducation*, 12(2-3), 17-27.
- Trahan, M., Carbonneau, M., Robillard, R., & Boisvert, M.-T. (1983). Interventions évaluatives : impact sur les évalués. *Mesure et évaluation en éducation*, 6(2), 69-78.
- Vial, M. (1995). Nature et fonction de l'auto-évaluation dans le dispositif de formation. *Revue française de pédagogie*, 112, 69-76.
- Vial, M. (2006). Les relations entre formation et évaluation : perspectives de recherches. *Mesure et évaluation en éducation*, 29(1), 81-97.
- Weiss, J. (1984). Heurs et malheurs d'un instrument d'évaluation. *Mesure et évaluation en éducation*, 7(1), 31-42.
- Weiss, J. (1994). Évaluer autrement ! *Mesure et évaluation en éducation*, 17