

Une resocialisation de l'évaluation des apprentissages

Joëlle Morrissette

Université Laval
joelle.morrissette@fse.ulaval.ca

Résumé

Cet article prend appui sur la problématisation développée dans le cadre d'une recherche doctorale au cours de laquelle j'ai réalisé un examen critique du domaine de l'évaluation des apprentissages afin d'éclairer l'émergence de l'évaluation dite formative aux côtés de la conception traditionnelle liée à la sanction des apprentissages. D'une part, cette réflexion met en relief qu'une vague de critiques relatives aux pratiques de classification a concouru à mettre en lumière des enjeux associés à leur emploi en contexte scolaire, notamment par le biais de l'évaluation qui sanctionne et certifie, suscitant un intérêt pour une évaluation susceptible de soutenir les apprentissages. D'autre part, réalisée dans un registre épistémologique, elle met également en relief l'intérêt d'une resocialisation de l'évaluation, c'est-à-dire d'une évaluation conçue telle une pratique qui procède de décisions s'appuyant sur des habitudes de pensée ainsi que sur des rapports de force entre divers groupes, et qui soulève des enjeux importants compte tenu que toutes pratiques de classification introduisent une norme à laquelle sont comparés les élèves.

Introduction

Il y a déjà plusieurs décennies que des chercheurs et chercheuses dans le domaine de l'éducation ont proposé l'idée d'une évaluation formative. Initialement avancée par Cronbach (1963) qui l'a utilisée pour désigner un processus de révision et de correction, l'idée a été reprise et reformulée par Scriven (1967) qui l'a employée pour désigner des moyens d'enseignement ou des méthodologies particulières d'évaluation des programmes d'études. Cependant, c'est surtout à la suite des travaux de Bloom, Hasting et Madaus (1971) qu'elle sera associée de façon plus précise à l'évaluation des apprentissages des élèves, et qu'un lien entre l'évaluation et une fonction de soutien aux apprentissages a été établi par certains membres de la communauté scientifique¹. Plusieurs experts soutiennent que l'évaluation formative est d'une grande importance en raison des possibilités d'amélioration des processus d'apprentissage qu'elle favoriserait (Black et Wiliam, 1998; Perrenoud, 1998; Scallon, 2000). Ils font aussi valoir que l'évaluation formative constitue un élément-clé de la différenciation pédagogique en permettant d'explorer des difficultés éprouvées par un élève et d'élaborer une démarche sur mesure visant à l'aider à les surmonter. Ainsi, l'évaluation formative serait au cœur de la réussite des élèves. Dans cette perspective, on comprend pourquoi l'évaluation formative est aujourd'hui l'objet de l'attention d'un grand nombre de chercheurs et de chercheuses. Cependant, c'est dans un recadrage sociohistorique qui s'attarde au contexte de son émergence, aux côtés de la conception traditionnelle de l'évaluation liée à la sanction des apprentissages, qu'on est plus à même de saisir les raisons du développement de cet intérêt.

La réflexion proposée ici consiste en un examen critique du domaine de l'évaluation des apprentissages réalisé dans une perspective interdisciplinaire (Larochelle et Désautels, 2006), examen qui prend appui sur la problématisation avancée dans le cadre d'une recherche doctorale (Morrissette, 2009).

Les pratiques de classification à l'école

Le développement d'un intérêt pour une évaluation formative repose sur plusieurs éléments contextuels, en particulier sur une vague de critiques qui ébranle le domaine de l'évaluation des apprentissages depuis les années 1970. À l'origine, ce domaine était alors défini presque entièrement par la fonction de certification et de contrôle de l'évaluation. En effet, le courant psychométrique américain, centré sur la mesure considérée scientifique (*scientific measurement*), marque la pensée en éducation en Amérique du Nord depuis la deuxième moitié du 20^e siècle. Tel que le relèvent Broad et Wade (1987), l'essor de ce courant a été en grande partie tributaire du contexte de la Deuxième Guerre mondiale et des rivalités entre les armées américaines et soviétiques, et sa dominance a tenu – et tient toujours – en grande partie au recours aux mathématiques dans les méthodes qui les mobilisent. Comme le souligne Lévy-Leblond (1984), plusieurs voient en la présence de mathématiques en recherche LE critère de scientificité. En effet, les mathématiques jouissent d'une réputation rarement remise en question: elles seraient neutres, pures et nécessaires. Cependant, pour Lévy-Leblond, les mathématiques sont bel et bien sociales, manipulées par des personnes qui ont des projets précis, et ne sont donc pas à l'abri de la «subjectivité» humaine. Cette position rejoint celle de Thuiller (1987), auteur de différents ouvrages sur l'épistémologie des sciences dans lesquels il met en relief l'inscription socioculturelle des sciences. Ce dernier fait valoir que la conviction que n'est science que ce qui est mesurable vient d'une tradition plus que millénaire et toujours très vivace au sein d'un «milieu culturel de plus en plus porté au réalisme, à l'étude exacte et positive des phénomènes» (Thuiller, 1987, p. 259-260). L'auteur rappelle que la science qui produit un savoir quantitatif se construit en privilégiant certaines catégories, les objets qui ne sont pas mesurables étant négligés, voire dévalorisés. Ainsi, sous l'influence forte du courant psychométrique américain et sous celle de la théorie comportementale behavioriste, promue notamment par les Watson, Hull, Skinner et Gagné, se développent une multitude de pratiques de classification telles que la mesure des comportements des élèves, de leurs caractéristiques dites psychologiques et de leurs apprentissages. Parmi ces pratiques, on trouve l'évaluation appelée sommative, le corollaire de la mesure des apprentissages associée à la sanction des études, qui intervient généralement au terme d'un processus d'enseignement / apprentissage ou de formation afin de « mesurer » les acquis des élèves dans une logique comptable, comme l'appellation l'indique.

Selon d'aucuns, l'évaluation sommative aurait fortement imprégné les pratiques scolaires au Québec (Kasadi, 2005) comme ailleurs (Black et Wiliam, 1998; Shepard, 2001). En effet, jusque vers la fin des années 1960, on considérait ces pratiques qui reposaient sur la mesure comme incontestablement «scientifiques»², et en conséquence, elles ont connu une grande popularité: tout semblait quantifiable par le biais de tests appropriés. Ainsi, sous l'influence des recherches effectuées dans le domaine de la psychologie et du paradigme qui les soutient, les différents acteurs qui participent à la situation éducative ont reconnu aux tests une propriété «objective» susceptible de remédier à ce que plusieurs concevaient comme la subjectivité et l'incohérence du jugement des enseignantes et des enseignants. À ce propos, le point de vue de Fourez, Englebert-Lecomte et Mathy (1997; voir aussi Broad et Wade, 1987) est instructif pour se distancier des idées reçues concernant le couple objectivité / subjectivité, lequel sert souvent d'argument d'autorité dans des débats entre chercheurs et chercheuses, et entre ces derniers et des populations profanes. En effet, les auteurs soutiennent que ces notions doivent se comprendre à l'aune de la perspective épistémologique adoptée. Par exemple, une vision positiviste postule l'existence d'une réalité indépendante des sujets. Dans cette optique, l'objectivité serait une attitude qui permettrait à un sujet de garder ses conceptions intactes, à l'abri de la déformation que pourraient engendrer des doctrines et des idéologies, comme si le jugement à poser sur une situation pouvait être exempt de préjugés, de sentiments ou de valeurs. En ce sens, l'objectivité s'oppose à la subjectivité que plusieurs voient comme une vision «contaminée» par un individu ou une collectivité. Par ailleurs, toujours selon les mêmes auteurs, une vision constructiviste postule plutôt que la «réalité» n'existe pas indépendamment des sujets qui la créent. Ainsi, on peut employer l'idée d'objectivité en fonction d'un projet précis et à partir de critères socialement stabilisés. Selon cette vision, tout discours est marqué par les intentions du sujet qui le structure et, par conséquent, objectivité et subjectivité se conçoivent comme étant des notions complémentaires.

Des critiques venant du domaine de l'évaluation

Des chercheurs et des chercheuses du domaine même de l'évaluation des apprentissages remettent en cause la valorisation de la mesure à partir des années 1970. En effet, plusieurs s'interrogent sur les pratiques d'évaluation et sur les prétentions associées à leur potentiel de scientificité. Par exemple, en 1978, Williams, Mosby et Hinson relatent les conclusions d'un grand nombre de rapports qui mettent au jour des biais culturels des tests employés pour classer les enfants autres que ceux appartenant à l'ethnie dominante, tels des Amérindiens, des Porto Ricains,

etc., et des erreurs d'interprétation qui sont commises pendant tout le processus de passation des tests et de leur analyse. Ces auteurs font valoir que la principale conséquence de l'emploi fréquent de ces tests est au final, l'exclusion de ces enfants du système scolaire. L'argumentation de ces auteurs soulève ainsi la question des intentions à l'origine de ces pratiques de *testing*, donnant à penser que certains chercheurs et chercheuses se servent de la notoriété de la mesure à des fins politiques ou économiques, ou encore pour cacher des intentions ségrégationnistes. D'autres, parmi lesquels figurent Hobbs (1975) ainsi que Crooks (1988), ce dernier ayant réalisé une recension abondamment citée dans le domaine de l'évaluation des apprentissages, déplorent l'impact des pratiques d'évaluation sommative sur l'estime de soi des élèves et, par conséquent, sur leur progression au sein d'un système scolaire. Enfin, d'autres encore (par exemple, Keeves et Alagumalai, 1998; Koretz, 1988; Linn, 2000) remettent en question la fiabilité des instruments employés pour mesurer les apprentissages, ébranlant ainsi la confiance qui leur était accordée jusqu'alors quant à leur capacité à les mesurer. Par exemple, dans leur ouvrage, Nichols et Berliner (2007) émettent de fortes réserves à l'endroit des tests ministériels quantitatifs. De fait, ces auteurs font valoir que la façon dont ces tests sont élaborés ainsi que l'orientation donnée à l'interprétation des résultats de ces tests sont fonction de projets politiques particuliers, et qu'il faut dès lors en faire un usage prudent en raison des impacts moraux négatifs sur les élèves et sur les enseignants et les enseignantes. Également, dès 1971, Mehan contribue à mettre en lumière que les tests employés à l'école pour mesurer les apprentissages des élèves font l'objet d'une forte réification qui repose sur une prétention à l'objectivité. En outre, ce chercheur fait valoir que ces pratiques de classification que constituent les tests employés à l'école impliquent des processus interprétatifs et donc interactionnels. Dans cette veine, d'aucuns parlent d'un risque d'une technologisation de l'évaluation, voire d'une «illusion instrumentale» (Bain, 1998; Scallon, 1988; Vial, 2001).

Des critiques venant d'autres domaines

Des travaux réalisés dans d'autres domaines permettent également de mettre en question les différentes pratiques de classification qui jouissent d'une grande popularité auprès de certains membres de la communauté scientifique. En effet, le caractère ambigu de l'évaluation et les risques liés aux usages des pratiques évaluatives sont également pointés du doigt par des chercheurs issus de différents domaines.

Le paléontologue Gould a publié en 1983 *La mal-mesure de l'homme*³, un livre qui attaque de front la question du classement des groupes humains par la mesure de leur « intelligence », en montrant notamment que la mesure soutient souvent l'idéologie ségrégationniste de chercheurs et de chercheuses qui se servent de tests dits objectifs pour soutenir leur thèse. Pour faire cette démonstration, il met en évidence les faiblesses des recherches qui reposent sur ce que plusieurs appelaient la science de la craniométrie au 19^e siècle et sur les tests d'intelligence très populaires au 20^e siècle, lesquels conduisaient inévitablement à désavantager certains groupes au sein d'une population. Par exemple, Gould (1983) rapporte le cas de Terman, Goddard et Yerkes, des chercheurs américains qui ont détourné de sa finalité première le test de QI conçu par Alfred Binet en s'en servant pour prouver l'infériorité ou la supériorité de certaines populations. En conséquence de leurs travaux, des jeunes Mexicains, Amérindiens et personnes à la peau noire ont été exclus du processus de scolarisation habituel et dirigés vers des classes spéciales où on offrait une formation plus manuelle. Ainsi, un des intérêts majeurs de l'ouvrage de Gould (1983) réside dans la mise en relation des conclusions de ces recherches avec le contexte sociohistorique dans lequel les scientifiques les élaborent, ce rapprochement lui permettant de mettre en relief que des préjugés, des joutes de pouvoir et même des considérations économiques conduisent souvent des chercheurs et chercheuses à avancer des conclusions biaisées. Ainsi, à l'instar de Fourez et *al.* (1997), Gould (1983) contribue à montrer qu'aucun instrument de mesure n'est neutre: la rigueur prétendue des procédés évaluatifs est toujours fonction d'un projet et d'une manière d'être et de penser; toute rigueur ou objectivité est partielle, contextualisée et relative à ce qui est visé.

Les publications de l'anthropologue Rist nourrissent également la mise en question des pratiques de classification. Pour élaborer sa critique, il s'appuie surtout sur les travaux de Becker (1963) sur la déviance, lesquels ont révolutionné la façon de voir la délinquance en attirant l'attention non pas sur les « propriétés internes » desdits déviants, mais notamment sur la société dans laquelle se trouvent ces déviants, et donc sur les conditions sociales, telles que les normes en vigueur, qui entourent l'attribution d'une telle étiquette. Les travaux de Becker proposent un changement radical de perspective qui amène à déplacer l'attention de l'élève vers ce qui se passe dans l'école, vers le processus par lequel on arrive à accoler une étiquette à un élève. Sur ces bases, Rist (1977) étudie la scolarisation dans des ghettos urbains, et conclut que les divers mécanismes d'évaluation mobilisés dans les écoles, tant en raison de la manière dont ces mécanismes sont encouragés et (re)produits qu'en raison de leurs effets sur les rapports interpersonnels entre les principaux acteurs concernés, ont des conséquences à long terme. Par exemple, l'auteur

releve les effets de marquage et d'étiquetage qui peuvent modifier les choix offerts aux élèves à l'intérieur du processus de scolarisation. Ce faisant, il établit des liens entre des causes possibles de la délinquance et la façon dont les institutions scolaires fixent des étiquettes sur les élèves par le biais des mécanismes de tri qu'opèrent les pratiques d'évaluation. De plus, toujours en s'appuyant sur les travaux de Becker (1963), Rist souligne l'importance de comprendre que ce sont ceux et celles qui détiennent du pouvoir dans une institution, et dans la société de manière plus générale, qui établissent la norme à laquelle les élèves sont comparés lorsqu'ils sont évalués.

Des critiques envers l'évaluation et envers les pratiques qui en relèvent viennent également du côté de la philosophie, en particulier de la *French Theory*, soit un cadre théorique très populaire aux États-Unis depuis le milieu des années 1970 qui sert notamment de socle pour de nouveaux genres académiques qui émergent à cette époque tels que les *cultural studies*, les *gender studies* ou encore les *multicultural studies*. Foucault est l'une des figures de proue⁴ de la *French Theory*. Son célèbre livre *Surveiller et punir: naissance de la prison*, paru en 1975, dénonce comment, dans les sociétés modernes, le pouvoir se manifeste dans les relations pédagogiques fondées sur une organisation minutieuse de la discipline. Foucault considère que l'école est une machine à contrôler et à normaliser les conduites – d'où l'idée de surveiller – puisqu'elle promeut une certaine manière de voir et d'agir par la mise en place de règles par exemple; en cela, soutient l'auteur, l'école s'assure du coup le pouvoir qui découle de ses mécanismes de contrôle. En effet, les pratiques qu'on associe aujourd'hui à l'évaluation qui sanctionne et certifie, telles que la notation ou d'autres façons de classer les élèves en fonction d'une norme, représenteraient de bons moyens d'assurer une domination, d'assurer un contrôle. En outre, dans son ouvrage, Foucault (1975) montre de façon convaincante comment l'examen constitue l'une de ces pratiques qui procurent à l'évaluateur un pouvoir de surveillance qui permet de qualifier, de classer, voire de punir, car l'examen compare les performances singulières à un ensemble qui représente la norme à suivre en traçant la limite qui définit la différence par rapport à toutes les différences. L'auteur affirme que l'examen réduit les élèves à des objets descriptibles, analysables et figés, leur attribuant des traits ou des propriétés qui sont vus comme leur étant inhérents; en d'autres mots, il établit sur les élèves une visibilité à travers laquelle on les différencie et on les sanctionne. Bref, la pensée de Foucault (1975) amène à concevoir que l'examen et d'autres types de pratiques de classification sont de puissants outils de normalisation.

Des critiques envers l'évaluation qui sanctionne et certifie les apprentissages viennent également du côté de la tradition sociologique, notamment de la plume de Perrenoud (1995) qui fait valoir que les pratiques qui s'inscrivent dans cette conception de l'évaluation provoquent le développement de comportements qui détournent l'élève de l'apprentissage. Selon l'auteur, ces pratiques engendreraient des sortes d'adaptations secondaires chez l'élève, c'est-à-dire des ajustements qui lui permettraient de survivre aux contraintes de sa situation obligée à l'école, à savoir à celles du système dans lequel il évolue. Perrenoud (1995) soutient que dans le contexte de l'évaluation traditionnelle qui sanctionne, l'élève tenterait de se préserver une part de liberté en cherchant à infléchir les règles du système, à faire illusion, à masquer ses « failles » et à mettre en valeur ses points forts. L'élève développerait ainsi des formes de résistances, des adaptations, telles que le plagiat, le bachotage ou encore des jeux de séduction pour négocier de l'aide ou des annotations plus favorables. Par ces propositions, le sociologue met en évidence un aspect souvent gommé de la situation de l'élève, à savoir qu'il exerce un métier qui n'est pas librement choisi et qui dépend constamment de l'évaluation qu'un tiers fait de ses qualités, de ses défauts, de son intelligence, de sa culture, etc. En ce sens, être un élève signifie être constamment sous le regard et le contrôle d'autrui; cela signifie également risquer étiquettes, sanctions et échecs, ce qui fait du métier d'élève un métier beaucoup plus contraignant que ne le sont la plupart des autres types d'occupation. Cet éclairage de Perrenoud (1995) contribue à mettre en lumière que les pratiques d'évaluation qui visent la sanction ou la certification des apprentissages ont des effets imprévus et non souhaités.

Toujours dans la tradition sociologique, Bourdieu et Passeron (1970) critiquent aussi les mécanismes de classification dont l'école fait usage en montrant notamment comment elle reproduit la structure de la distribution du capital culturel. Sur la base de nombreuses recherches empiriques, ces auteurs proposent un cadre théorique autour de ce qu'ils conçoivent comme des structures de domination symbolique au sein desquelles le système scolaire occupe une place de choix. En effet, pour eux, l'école est une institution qui fonctionne comme une machine de sélection sociale; par différents mécanismes de classification et de sélection opérés en majeure partie par les pratiques d'évaluation des apprentissages, elle reproduirait les disparités sociales et les conditionnements; elle sanctionnerait ainsi les élèves qui ont une culture différente – et donc qui appartiennent à une classe sociale distincte – de celle qu'elle valorise, s'attribuant un pouvoir discrétionnaire tributaire d'un idéal culturel arbitraire auquel sont comparés les élèves. Plus que Foucault (1975) qui reconnaît que les acteurs ont une réflexivité leur permettant d'agir

et ne sont pas que les rouages du système, Bourdieu et Passeron (1970) attribuent au pouvoir et aux relations auxquelles il est associé un caractère déterminant⁵. Troger (2002) relève que l'analyse réalisée par ces sociologues du rôle de l'école dans la reproduction des inégalités sociales a significativement influencé les recherches ultérieures en sociologie de l'éducation, et même certaines produites dans d'autres disciplines telles qu'en histoire.

La resocialisation des pratiques de classification

La vague de critiques adressées aux pratiques de classification et à leur usage plus particulier en contexte scolaire concourt à identifier des enjeux et des risques liés aux pratiques évaluatives associées à la sanction et à la certification des apprentissages. Il s'agit là d'un regard qui *resocialise* l'évaluation, selon le mot de Larochelle (2003), qui invite à considérer que les pratiques qui sont réalisées en son nom organisent les manières de voir à travers un travail de production de catégories. Aux côtés de Monceau (2001) et de Becker (1963), Larochelle propose d'envisager:

la classification comme le porte-parole de ceux et celles qui l'ont inventée et de prendre acte ainsi que, dans son processus comme dans ses produits, elle comporte matière à réflexion socioépistémologique: la classification procède de décisions, est rendue possible par des habitudes et des patterns de pensée, de parler et, enfin, participe d'enjeux et de tensions non négligeables compte tenu que toute classification revient à instaurer un passage obligé quant à ce qui doit être tenu pour légitime au sein d'une société ou d'un groupe donnés (p. 59-60).

Ainsi, d'une part, Larochelle (2003) fait valoir que les pratiques de classification renseignent davantage sur les projets de leurs auteurs que sur les choses ou les personnes classées; la chercheuse ajoute que les utilisateurs de ces classifications oublient souvent ces projets dès lors qu'elles trouvent différents usages politiques et scientifiques, voire économiques. D'autre part, Larochelle soutient que les classifications procèdent d'habitudes de pensée tellement imbriquées dans des pratiques quotidiennes qu'on finit par ne plus envisager que d'autres perspectives sont possibles. En ce sens, Bourdieu (1972) dirait que ces pratiques relèvent de l'habitus, un concept abondamment utilisé en sociologie qu'il définit comme un « lot de schèmes permettant d'engendrer une infinité de pratiques adaptées à des situations toujours renouvelées, sans jamais se constituer en principes explicites » (p. 209).

La resocialisation de l'évaluation des apprentissages invite d'ailleurs à envisager les pratiques scolaires évaluatives comme faisant partie d'une mécanique institutionnelle résultant de rapports de force entre les individus et / ou les groupes sociaux qui sont impliqués de près ou de loin au système scolaire. En fait, l'évaluation représente et sanctionne la conformité à la culture scolaire qui serait celle de la culture dominante. En ce sens, l'évaluation serait porteuse d'un pouvoir qui influence la situation éducative en établissant des normes. Il convient tout de même de nuancer en rappelant, comme le note Chervel (1998), que la culture scolaire ne se réduit pas au *cahier des charges*, c'est-à-dire au programme d'étude officiel que les différents mécanismes d'évaluation doivent sanctionner. En effet, en parlant de *l'effet Wilamowitz*, l'auteur fait valoir qu'il y a un ensemble « d'effets culturels non prévisibles, engendrés par le système scolaire en toute indépendance » (p. 191). En appui à son propos, il donne l'exemple de la formation des maîtres français à l'orthographe entre les années 1830 et 1850 qui a donné lieu à plusieurs effets imprévus, dont l'arrêt de l'évolution de l'orthographe qui, jusque-là, s'était adaptée progressivement aux changements phonétiques qui s'opéraient dans la langue. En ce sens, les pratiques d'évaluation des apprentissages ne reconduiraient pas systématiquement la culture dominante.

Conclusion

Cet examen critique du domaine de l'évaluation des apprentissages permet, d'une part, d'attirer l'attention sur les enjeux associés aux usages de l'évaluation en contexte scolaire, des enjeux liés, comme on l'a vu, à l'estime de soi des élèves, à leur cheminement dans le processus de scolarisation, à la reproduction d'une certaine hiérarchie sociale, etc. Compte tenu de la force symbolique des pratiques d'évaluation dans nos sociétés, qui repose sur des croyances à l'égard des instruments de mesure qui confèreraient l'assurance d'une scientificité par leur supposé potentiel objectif et apolitique, il importe d'adopter une perspective resocialisante de l'évaluation des apprentissages pour éviter d'occulter les rapports sociaux – dont les relations de pouvoir – qui sont inévitablement parties prenantes de toute classification.

D'autre part, cet examen critique permet de comprendre pourquoi, depuis le début des années 1970, plusieurs chercheurs et chercheuses se sont intéressés à une conception différente de l'évaluation qui se centrerait sur l'idée de favoriser les apprentissages plutôt que de les sanctionner ou de les certifier. En effet, l'évaluation formative des apprentissages viserait à servir l'apprentissage et l'enseignement, et serait centrée sur les processus, et non sur les produits de l'apprentissage. En raison du contexte de réformes éducatives dans différents pays occidentaux visant à amener un plus grand nombre d'élèves à franchir avec succès le processus de scolarisation, l'évaluation formative est d'ailleurs au cœur des préoccupations actuelles; en témoigne la publication abondante de travaux de recherche sur le sujet⁶ et la mise en place d'études internationales telles que celle du Centre pour la Recherche et l'Innovation dans l'enseignement (2005) menée de façon concomitante dans différents pays. Il y a donc un intérêt certain à documenter les manières par lesquelles les enseignants et les enseignantes mettent en œuvre l'évaluation formative, notamment pour comprendre les enjeux dont elles sont porteuses, mais aussi pour comprendre le système de pratiques dans lequel elles s'inscrivent. En effet, comment ces praticiens et ces praticiennes composent-ils avec l'évaluation des apprentissages de leurs élèves au quotidien? Plus particulièrement, comment négocient-ils la tension entre la finalité de soutien aux apprentissages de l'évaluation formative et la forte pression sociale actuelle en faveur d'un système éducatif plus performant, ce que d'aucuns nomment l'obligation de résultats (par exemple, Lessard, 2004; Meirieu, 2004)?

Bibliographie

- Bain, D. (1998). L'évaluation formative fait fausse route. *Mesure et Évaluation en Éducation*, 10, 23-32.
- Becker, H.S. (1963). *Outsiders: Studies in the sociology of deviance*. New York: Free Press.
- Black, P. (1986). Assessment for learning. In D.L. Nuttall (Ed.), *Assessing educational achievement* (p. 7-18). London: Falmer Press.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.
- Bloom, B.S., Hasting J.T. & Madaus, G.F. (1971). *Handbook on formative and summative evaluation of student learning*. New York: Mc Graw-Hill.
- Boudon, R. (1973). *L'inégalité des chances*. Paris: Armand Colin.
- Bourdieu, P. (1972). *Esquisse d'une théorie de la pratique*. Genève: Droz.
- Bourdieu, P. & Passeron, J.-C. (1970). *La reproduction: éléments pour une théorie du système d'enseignement*. Paris: Minuit.
- Broad, W. & Wade, N. (1987). *La souris truquée. Enquête sur la fraude scientifique* (C. Jeanmougin, trad.). Paris: Seuil.
- Centre pour la Recherche et l'Innovation dans l'enseignement. (2005). *L'évaluation formative: pour un meilleur apprentissage dans les classes secondaires*. Paris: Organisation de coopération et de développement économiques.
- Chervel, A. (1998). *La culture scolaire: une approche historique*. Paris: Belin.
- Cronbach, L.J. (1963). Course improvement through evaluation. *Teachers College Record*, 64(5), 672-683.
- Crooks, T.J. (1988). The impact of classroom evaluation practice on students. *Review of Educational Research*, 58, 438-481.
- Foucault, M. (1975). *Surveiller et punir: naissance de la prison*. Paris: Gallimard.
- Fourez, G., Englebert-Lecomte, V. & Mathy, P. (1997). *Nos savoirs sur nos savoirs. Un lexique d'épistémologie pour l'enseignement*. Bruxelles: De Boeck.
- Gould, J.S. (1983). *La mal-mesure de l'homme*. Paris: Ramsay.
- Hobbs, N. (Eds.). (1975). *Issues in the classification of children* (vol. 1). San Francisco: Jossey-Bass.
- Kasadi, C. (2005). *Exploration des pratiques de professeurs des mathématiques du secondaire à l'égard de l'évaluation formative en mathématiques*. Thèse de doctorat non publiée, Université du Québec à Trois-Rivières en association avec l'Université du Québec à Montréal, Trois-Rivières, QC.
- Keeves, J. & Alagumalai, S. (1998). Advances in measurement in science education. In B.J. Fraser & K.B. Tobin (Eds.), *International handbook of science education* (p. 1229-1244). Dordrecht, NL: Kluwer Academic Publishers.
- Koretz, D. (1988). Arriving at Lake Wobegon: Are standardized tests exaggerating achievement and distorting instruction? *American Educator*, 12(2), 8-15, 46-52.
- Larochelle, M. (2003). Rapport au savoir et socialisation à la cité scientifique. In S. Maury & M. Caillot (dir.), *Rapport au savoir et didactiques* (p. 51-79). Paris: Fabert.
- Larochelle, M. & Désautels, J. (2006). L'éducation aux sciences et le croisement des expertises. In A. Legardez & L. Simonneaux (dir.), *L'école à l'épreuve de l'actualité* (p. 61-79). Paris: ESF.
- Lessard, C. (2004). L'obligation de résultats en éducation: de quoi s'agit-il? Le contexte québécois d'une demande sociale, une rhétorique du changement et une extension de la recherche. In C. Lessard & Ph. Meirieu (dir.), *L'obligation de résultats en éducation* (p. 23-48). Québec, QC: Presses de l'Université Laval.
- Lévy-Leblond, J.-M. (1984). *L'esprit de sel. Science, Culture, Politique*. Paris: Seuil.
- Linn, R. (2000). Assessments and accountability. *Educational Researcher*, 29, 4-16.
- Mehan, H.B. (1971). *Accomplishing understanding in educational settings*. Unpublished dissertation, University of California, Santa Barbara, CA.
- Meirieu, Ph. (2004). L'école entre la pression consumériste et l'irresponsabilité sociale. In C. Lessard & Ph. Meirieu (dir.), *L'obligation de résultats en éducation* (p.5-21). Québec, QC: Presses de l'Université Laval.
- Monceau, G. (2001). De la classification des individus à celle de leurs devenir dans l'institution scolaire. *La Lettre du Groupe de Recherche et d'Action pour l'Enfance et l'Adolescence (GRAEA)*, 43, 27-35.
- Morrisette, J. (2009). *Analyse des manières de faire l'évaluation formative des apprentissages explicités par un groupe d'enseignantes du primaire: une perspective interactionniste*. Thèse de doctorat non publiée, Université Laval, Québec, QC.

- Nichols, S.L. et Berliner, D.C. (2007). *Collateral damage. How high-stakes testing corrupts America's schools*. Cambridge, MA: Harvard Educational Press.
- Perrenoud, Ph. (1998). *L'évaluation des élèves. De la fabrication de l'excellence à la régulation des apprentissages*. Bruxelles: De Boeck.
- Perrenoud, Ph. (1995). *Métier d'élève et sens du travail scolaire* (2e édition). Paris: ESF.
- Rist, R.C. (1977). On understanding the processes of schooling: The contributions of labelling theory. In J. Karabel & A.H. Halsey (Eds.), *Power and ideology in education* (p. 292-305). New York: Oxford University Press.
- Scallon, G. (2000). *L'évaluation formative*. Québec, QC: Renouveau pédagogique.
- Scallon, G. (1988). *L'évaluation formative des apprentissages: la réflexion*. Québec, QC: Presses de l'Université Laval.
- Scriven, M. (1967). The methodology of evaluation. In R.W. Tyler, M. Gagné & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (p. 39-83). Chicago: Rand McNally.
- Shepard, L.A. (2001). The role of classroom assessment in teaching and learning. In V. Richardson (Ed.), *Handbook of research on teaching* (4th edition) (p. 1066-1101). Washington, DC: American Educational Research Association.
- Thuillier, P. (1987). À propos des fondements «symboliques» de la science moderne: le livre, la loi, le nombre, la machine, le primat du mâle. *Cahiers de la Fondation Archives Jean Piaget*, 8, 255-264.
- Troger, V. (2002). Bourdieu et l'école: la démocratisation désenchantée. *Sciences Humaines*, numéro spécial, 1, 16-23.
- Vial, M. (2001). *Se former pour évaluer. Se donner une problématique et élaborer des concepts*. Bruxelles: De Boeck.
- Williams, R.L., Mosby, D. & Hinson, V. (1978). Critical issues in achievement testing of children from diverse ethnic backgrounds. In M. Wargo & D.R. Green (Eds.), *Achievement testing of disadvantaged and minority students for educational program evaluation* (p. 41-72). Monterey, CA: CTB & McGraw Hill

Notes

¹ Ici, il est question de l'évaluation formative en tant qu'objet reconnu par les chercheurs et chercheuses; tel que le soutient Black (1986), il est raisonnable de penser que, dans la mise en œuvre de leurs pratiques, les enseignantes et les enseignants ont toujours fait une forme d'évaluation pour soutenir les apprentissages des élèves, et ce, même si l'appellation n'était pas encore stabilisée dans la communauté scientifique.

² Selon Fourez, Englebert-Lecomte et Mathy (1997), qui ont publié un lexique d'épistémologie pour l'enseignement, est dit «scientifique» ce qui est produit par une communauté scientifique, reconnu par les membres de cette communauté et aussi, le plus souvent, par d'autres groupes sociaux. Ainsi, selon ces auteurs, ce qualificatif implique notamment la reconnaissance sociale dont quelque chose fait l'objet et son inscription dans un paradigme.

³ Intitulé *Mismeasure of Man* (1981) dans sa version originale de langue anglaise, l'ouvrage de Gould reçoit en 1982 le prix du *National Book Critics' Circle*.

⁴ Parmi les autres figures marquantes de la *French Theory*, on trouve les philosophes français Derrida et Deleuze, ainsi que le sociologue Bourdieu.

⁵ Plusieurs critiquent la dimension déterministe de l'œuvre de Bourdieu. Boudon (1973), par exemple, lui reproche de brosser un tableau de l'école exclusivement articulé autour des structures déterminantes du système qui ne laissent aucune place à la réflexivité des acteurs, à leur marge de manœuvre. Toutefois, ce déterminisme est plus nuancé dans les dernières publications de Bourdieu à qui on doit reconnaître, malgré les critiques qu'on peut lui adresser, un héritage intellectuel marquant.

⁶ Voir la recension d'écrits réalisée dans le cadre de la recherche doctorale (Morrissette, 2009).