



Tâches pour apprendre, tâches pour évaluer

Lucie MOTTIER LOPEZ¹ (Université de Genève, Suisse)

Les mots, c'est l'enfer pour Bronckart (1987) : « les mots nous tiennent captifs, plutôt que de nous montrer le réel, ils nous empêchent de le voir et nous désorientent ... Les mots, c'est les autres, ça vient d'ailleurs et ça nous enferme » (pp. 37-38). Une *tâche* est-elle synonyme de *problème* ? Tout *problème* est-il un *exercice* ? Un *devoir* est-il du même ordre qu'un *exercice* et une *tâche* ?... Quels sens donnons-nous à ces mots omniprésents dans le monde éducatif ?

Ainsi, entendra-t-on dans la bouche d'un élève : « On a fait des *exercices* de maths ce matin » ; « Le prof nous a donné des *problèmes* difficiles dans son éval ». Ou de la bouche des parents : « As-tu fait tes *devoirs* pour demain ? ». Les usagers de l'école (enseignant.e.s, élèves, parents, etc.) partagent un sens commun de ces mots. Certains appartiennent plus spécifiquement au langage de la profession enseignante. Dans le canton de Genève par exemple, conformément aux prescriptions institutionnelles en matière d'évaluation certificative à l'école primaire, les enseignant.e.s décrivent les *exercices* de leurs activités-bilans en termes de *tâches* complexes (le plus souvent associées à des énoncés de *problèmes*) et de *tâches* spécifiques (quand il s'agit par exemple de mémoriser une table de multiplication, d'appliquer une règle, de résoudre un calcul lacunaire). Certains mots sont associés à des approches pédagogiques, telle la pédagogie par *situation-problème* ou les approches par compétences qui accordent une place importante aux notions de *tâche* et de *situation* (entre autres). Encore plus généralement, certains de ces termes s'inscrivent dans des champs conceptuels propres à des cadres théoriques donnés – par exemple les notions de *problème* et d'obstacle, ou de *tâche*, *situation*, milieu (etc.) dans les théories didactiques.

Les contextes d'usage se croisent, s'interpénètrent : un savoir scientifique approprié peut devenir une connaissance professionnelle qui, peu à peu, devient un langage en partie partagé avec les usagers de l'école. Les significations se transforment, les attributs définitoires s'estompent. Ce qui, à l'origine était (peut-être) un concept, devient un sens commun et s'entremêle, voire se confond à des mots déjà existants. La tâche devient l'exercice, l'exercice est la tâche... Le chercheur lui-même doute, s'interroge, tente d'élaborer des sens circonscrits, conscient que souvent ses définitions ne sont que partiellement distinctes du sens commun. Ceci dit sans critique aucune. Sinon, cela signifierait que les différents « mondes » de la

1. Contact : Lucie.Mottier@unige.ch



recherche et de la pratique, comme parfois on aime les appeler, ne communiqueraient pas, ne se rencontreraient pas, ne « s'inter-féconderaient » pas et c'est alors cela qui serait critiquable.

En préambule

Pour ma part, je considère que les tâches, exercices, devoirs ont pour point commun d'être des dispositifs d'enseignement en tant que *moyens-buts* (au sens de Dewey, 2011) de médiation pour soutenir l'apprentissage de l'élève, un apprentissage qui est à la fois individuel et foncièrement social. Certains n'hésitent d'ailleurs pas à qualifier cet apprentissage de « pratique sociale spécifique » associée à des communautés culturelles qui lui donnent forme (e.g., Lave & Wenger, 1991). Dans ce cas, toute tâche, exercice, devoir suscitant l'activité de l'apprenant (et d'un collectif) se comprend au regard d'un cadre social et culturel immédiat et plus large, porteur de finalités, d'usages, de significations vues comme partagées, préexistantes et émergentes (e.g., Cobb et al., 1997). Je ne considère cependant pas ces termes comme des synonymes. J'aurais tendance à penser que toute tâche n'est pas un problème (qui implique une certaine démarche intellectuelle).² Mais que tout problème, devoir, exercice proposé aux élèves impliquent la réalisation d'une ou plusieurs tâches, en tant que composantes d'une situation plus large marquée par des contrats, des normes, des pratiques, des contextes socio-institutionnels.

Une abondante littérature scientifique, internationale, porte sur la notion de tâche dans le contexte scolaire : des tâches dont l'enjeu est l'apprentissage de l'élève, avec notamment une centration des approches psychologiques sur les processus cognitifs impliqués dans la réalisation de la tâche (Le Not, 1999³) ; des tâches, dans une perspective didactique, qui sont conceptualisées comme étant des outils pour enseigner et qui matérialisent les objets enseignés (Schneuwly, 2001). Dans les travaux anglophones, initialement orientés plutôt vers l'activité psychologique de l'élève, des contributions actuelles cherchent aussi à mieux conceptualiser la relation entre les tâches dans des disciplines scolaires données, les « actions » de l'enseignant.e et la nature des apprentissages réalisés (e.g., Sullivan, Clarke, Clarke & O'Shea, 2010). Malgré les orientations théoriques différentes, un consensus semble apparaître, y compris dans les écrits anglophones, à propos de la distinction à faire entre la tâche qui est de l'ordre du prescrit, du déjà-là, et l'activité de la personne qui exécute la tâche, dans la lignée des propositions de l'approche ergonomique de Leplat et Hoc (1983). Il est peut-être utile de souligner aussi, en préambule, que la tâche ne suffit évidemment pas pour enseigner et apprendre sans les systèmes complexes de médiation, de régulation, de construction de sens qu'elle sollicite et les proces-

2. Il n'y a pas de consensus, certains considèrent que la tâche « consiste en un problème à résoudre » (e.g., Dolz, Schneuwly, Thévenaz & Wirthner, 2001, p. 5) et d'autres, tel Brousseau, font la différence entre une tâche en tant que « succession définie d'actions connues », tel un algorithme et un problème qui n'est pas une tâche.

3. <http://gallica.bnf.fr/ark:/12148/bpt6k1200503m>



sus de validation, d'institutionnalisation, d'évaluation qui lui sont associés. Se pose alors la question délicate de l'orientation de l'activité du sujet par le « prescripteur » voire concepteur de la tâche, l'enseignant.e dans le contexte scolaire, entre une absence de guidage versus un guidage trop fort débouchant sur des actions qui pourraient être inadéquates par rapport aux enjeux de la tâche.

Au regard de ces quelques éléments préalables, qu'en est-il plus précisément de la relation entre tâche et évaluation des apprentissages, une évaluation formative qui vise à soutenir la régulation de l'apprentissage de l'élève par des processus de guidage notamment, et une évaluation sommative et certificative qui vise le contrôle et la validation des acquis des élèves? L'article expose quelques éléments de réflexion en distinguant deux grands cas de figure, (1) quand l'évaluation est intégrée à une tâche d'apprentissage, (2) lorsque la tâche a une fonction d'évaluation des apprentissages réalisés. L'article se termine par quelques considérations sur les notions de situation et de contexte quand il s'agit, dans ces différentes tâches, de construire du sens et d'attribuer une « valeur » au travail de l'élève.

Des évaluations intégrées à la tâche d'apprentissage

Dans ce premier cas de figure, la tâche d'apprentissage – en tant qu'outil d'enseignement – est première par rapport à l'évaluation. Des processus d'évaluation, de différents ordres s'y intègrent, souvent qualifiés d'évaluation « continue ». Ils peuvent avoir lieu en cours de réalisation de tâche, à l'occasion d'échanges dialogiques entre l'enseignant.e et l'élève/ les élèves, ou de façon plus formelle par le moyen d'outils d'autoévaluation et autres « guides » critériés par exemple. Ils existent également en fin de tâche quand il s'agit de valider la production⁴ réalisée par l'élève.

Toute évaluation continue n'est cependant pas d'emblée formative, confusion que l'on peut parfois constater dans la littérature. L'évaluation peut être seulement *informative*, en tant que rétroaction fournie à l'élève sur son activité au regard des attentes et exigences (prédéterminées ou susceptibles de se moduler en fonction des productions effectives des élèves). Dans les échanges dialogiques entre enseignant.e et élève/s, l'évaluation se manifeste par un énoncé oral (appelé *feedback* dans une large littérature) pouvant porter sur différentes dimensions de l'activité de l'élève, dont non seulement sur les contenus de savoir en jeu et la réussite de la tâche, mais également sur la façon de la réaliser, les démarches d'autorégulation pour y parvenir, les caractéristiques personnelles de l'élève (catégories dégagées par la méta-analyse sur le *feedback* de Hattie & Timperley, 2007). En fin de tâche, cette évaluation informative relève d'une *validation* de conformité ou non conformité par rapport à l'attendu – à une « correction » de l'exercice par exemple, dans le langage de la classe mais sans autre intention formative par rapport à l'exploitation de la réponse de l'élève/des élèves.

4. La notion de production englobe ici l'ensemble des différentes formes de « réponse » ou « produit » proposées par l'élève.



Ces feedback (oraux/écrits, individuels/collectifs) sont constants dans les discours et pratiques de la classe. Ils donnent des repères aux élèves et à l'enseignant.e sur le « vrai/faux » dont Margolinas (1993) signale le rôle et l'importance dans les processus d'enseignement. Ils peuvent aussi être contreproductifs, comme montré dans les recherches citées par Hattie et Timperley ; ils peuvent faire obstacle à la compréhension de l'élève, provoquer des conduites d'évitement, par exemple.

Pour qu'elle soit réellement *formative*, l'évaluation intégrée à la tâche implique nécessairement des processus de médiation, de guidage, de construction conjointe de sens à des fins de régulation de l'enseignement et/ou de soutien à l'autorégulation de l'élève/des élèves. Autrement dit, le seul feedback n'est pas suffisant. En fin de tâche, les productions des élèves représentent des sources d'information pour envisager de nouvelles tâches adaptées, de nouveaux « exercices ». Parfois, cette évaluation formative se formalise par le moyen d'outils « physiques » sous forme de fiches d'autoévaluation, de guides, de grilles critériées à renseigner, etc. Ce faisant, un des enjeux est l'articulation qui demande à être pensée entre tâche de la discipline scolaire et tâche qui consiste, pour l'élève, à autoévaluer sa démarche, ses stratégies, sa production, etc. : une articulation dont certaines études ont montré les bénéfices tout comme les possibles écueils (e.g., Allal & Michel, 1993). Cette articulation dans une même situation, ou organisée dans une même séquence d'enseignement, entre les tâches disciplinaires (orientées notamment par les modèles des savoirs qui s'y concrétisent) et les tâches (auto)évaluatives (à des fins d'autorégulation) destinées à l'élève/au collectif représente un enjeu de recherche insuffisamment problématisé à mon avis. De mon point de vue, elle incite à conceptualiser l'évaluation à l'interface entre théories didactiques et théories sur les processus d'apprentissage et de régulation des élèves en contexte scolaire.

Tâches pour évaluer des contenus disciplinaires

Comme on vient de le voir, la tâche d'apprentissage peut comporter en son sein des processus évaluatifs de différents ordres (informatif, de contrôle, de validation, formatif, d'autoévaluation à des fins de régulation et de différenciation⁵). Mais la tâche peut aussi avoir une fonction première d'évaluation des acquis des élèves dans les situations d'évaluation sommative et certificative⁶.

5. Par manque de place, je ne développerai pas les nuances qui seraient encore à faire autour de l'idée d'évaluation-régulation dans une approche plutôt de gestion des apprentissages ou en tant que problématique du sens (voir Vial, 2012, par exemple).

6. Nous faisons une différence entre l'évaluation qui consiste à transmettre dans un document officiel, public et permanent une appréciation globale mais dont la visée n'est pas une certification formelle – comme par exemple dans certains bulletins scolaires qui demandent une appréciation des « apprentissages dans la vie scolaire », e.g. « respect des règles de la classe » – et une évaluation qui atteste des niveaux d'acquisition en rapport à des exigences curriculaires délimitées.



Il n'est pas inutile de rappeler que dans les travaux des premières heures associés au courant docimologique⁷, la conceptualisation de l'évaluation sommative s'appuyait sur un cadre théorique attaché au paradigme de la mesure pour penser l'élaboration d'« items » (questions ou parties de tests). On ne parlait pas ici de tâche. Une des préoccupations des chercheurs était de limiter la subjectivité des évaluateurs et contrôler les facteurs qui font varier leurs jugements de façon arbitraire. Pour ce faire, des méthodologies de construction de tests ont notamment été élaborées, recommandant certains items au regard de leurs propriétés métriques garantissant une meilleure validité des tests (e.g., Laveault & Grégoire, 2014). Des typologies de formats des items sont proposées autour de la distinction bien connue entre « questions ouvertes » (e.g., questions demandant une performance, questions à réponse narrative) versus « questions fermées » (e.g., questions à choix multiples, questions vrai-faux, questions d'appariement), et « réponse à produire » versus « réponse à choisir ». Des travaux sophistiqués au plan statistique sont aujourd'hui proposés, notamment la théorie des réponses aux items et des extensions de la théorie de la généralisabilité dans des approches éducatrice et diagnostique. Ici, on retiendra que le cadre de conceptualisation des items-questions-énoncés prescriptifs adressés à l'élève à des fins d'évaluation relève d'une logique qui n'est pas directement celle associée à l'enseignement et à l'apprentissage, contrairement aux processus évaluatifs intégrés aux tâches présentés plus haut.

Avec l'élargissement du modèle de l'évaluation formative dans le monde francophone (Allal & Mottier Lopez, 2005), l'émergence de modèles alternatifs tels l'évaluation dite authentique (*educative assessment*, Wiggins, 1988) ou l'évaluation-soutien d'apprentissage (*assessment for learning*, Black, Harrison, Lee, Marshall & Wiliam, 2003) dans le monde anglophone, la conception de l'évaluation-mesure est largement remise en question lorsqu'il s'agit de penser l'évaluation en classe (*classroom assessment*), y compris quand elle est sommative et certificative. Une réflexion s'engage alors sur la nature des tâches proposées aux élèves dans les évaluations. À titre d'exemple, nous résumons ci-dessous trois contributions d'auteurs qui donnent un aperçu – partiel – de différentes orientations des travaux.

Commençons par Wiggins, chercheur américain qui, à la fin des années 1980, propose une réflexion critique entre les « tests typiques » et ce qu'il appelle des « tâches authentiques ». Le tableau 1 reproduit les points principaux de comparaison.

Un ensemble de voix émettra des critiques face à cette notion d'authenticité. Pourtant, cette dernière sera largement reprise, y compris pour conceptualiser les situations d'apprentissage en classe. Elle aura un impact certain pour penser autrement les « questions » (au sens large) formulées dans les évaluations et leur validité. C'est la « logique de l'apprentissage » (ou logique d'évaluation « éducative » dans les termes de l'auteur) qui prévaut ici pour conceptualiser et argumenter les choix évaluatifs en termes de tâches.

7. C'est-à-dire tout au long de la première partie du 20^e siècle, débouchant sur la publication dans les années 1960-1970 de bilans des travaux menés. Par exemple Piéron (1963).



Tableau 1 : Differences between typical tests and authentic tasks (Wiggins, 1988, chapitre 2, section 2, figure 2.1⁸)

Typical tests	Authentic tasks
Require correct responses only	Require quality product and / or performance, and justification
Must be unknown in advance to ensure validity	Are known as much as possible in advance; involve excelling at predictable demanding and core tasks; are note 'gotcha!' experiences
Are disconnected from a realistic context and realistic constraints	Require real-world use of knowledge: the student must 'do' history, science, etc. in realistic simulations or actual use
Contain isolated items requiring use or recognition of known answers or skills	Are integrated challenges in which knowledge and judgment must be innovatively used to fashion a quality product or performance
Are simplified so as to be easy to score reliably	Involve complex and non-arbitrary tasks, criteria, and standards
Are one shore	Are iterative; contain recurring essential tasks, genres and standards
Depend on highly technical correlations	Provide direct evidence, involving tasks that have been validates against core adult roles and discipline-based challenges
Provide a score	Provide usable, diagnostic (sometimes concurrent) feedback: the student is able to confirm results and self-adjust as needed

Le deuxième exemple est issu des travaux francophones. Nous avons choisi les travaux de Vial qui, spécialiste de l'évaluation, argumente dans les années 1990 en faveur d'une analyse didactique des tâches à partir de la typologie bien connue de Bloom: analyse, synthèse, évaluation. Dans sa contribution de 1996, l'auteur distingue deux grands ensembles de tâches :

- Des tâches à focalisation procédurale, associées à « des techniques dont on attend la maîtrise, l'intégration de codes » (p. 187) ;
- Des tâches à focalisation processuelle. Vial prend l'exemple du français et les associe à « des objets de construction de sens personnels comme l'écriture de poèmes, ... plus particulièrement ces tâches d'écriture où est manipulé l'espace symbolique du langage ... des tâches de créativité » (pp. 187-188).

Un ensemble de paramètres d'analyse sont délimités pour caractériser les tâches procédurales: nombre de procédures (tâche simple, tâche complexe), types de tâche (notionnelle, classe de problème, situation-problème), rôle (tâche d'apprentissage, tâche de vérification des acquis), nature (analyse, synthèse, contrôle). Ainsi, par exemple, il distingue :

8. Référence APA pour ebook.



- la tâche qui est conçue comme un moyen de «faire» apprendre de nouveaux savoirs ;
- la tâche qui sert à utiliser les savoirs appris comme des outils pour résoudre cette même tâche ;
- la tâche qui sert à entraîner des savoirs appris.

Une réflexion est menée sur les conséquences pour l'évaluation, notamment du point de vue des référents et des modèles de pensée (déterministe, structuraliste, cybernétique, systémiste) qui fondent le sens de différents «types» d'évaluation. De la démonstration de l'auteur, on retiendra qu'il plaide en faveur d'une «évaluation⁹ des tâches» susceptible de se mettre au service de «l'acquisition didactique» dans ce qu'il appelle une «herméneutique complexe». Le projet de l'auteur est de ne plus opposer les approches didactiques et modèles de pensée propres à l'évaluation des apprentissages.

Enfin, le troisième exemple est caractéristique des approches par compétences qui ont émergé dans le monde de l'école courant des années 2000 et du défi¹⁰ de pouvoir évaluer des compétences. La contribution de Rey, Carette, Defrance et Kahn (2003) donne à voir la conceptualisation de la relation vue comme constitutive entre la tâche et ses caractéristiques et la nature de l'activité cognitive sollicitée. Une «épreuve» d'évaluation diagnostique en trois phases est proposée. Elle a pour but d'identifier l'origine des difficultés des élèves : «Phase 1 : on propose aux élèves une tâche complexe exigeant le choix et la combinaison d'un nombre significatif de procédures que les élèves doivent posséder (...) Phase 2 : on propose la même tâche complexe, mais découpée en tâches élémentaires présentées dans l'ordre où elles doivent être accomplies pour réaliser la tâche globale (...) Phase 3 : on demande aux élèves d'accomplir les procédures exigées dans les phases précédentes, mais sous une forme décontextualisée (...)» (p. 92). Trois «degrés» de compétence sont associés à chaque phase, (1) compétence complexe, (2) compétence élémentaire avec interprétation (ou cadrage) de la situation, (3) compétence élémentaire ou procédure.

Il ressort notamment de cette modélisation que la définition de l'apprentissage évalué – ici, une compétence mais cela peut être pour tout type d'apprentissage – exige une mise en relation entre la nature de l'activité de l'élève/ des élèves et les caractéristiques des tâches. Concrètement, cela demande à l'évaluateur de définir «l'objet évalué», souvent formulé en termes d'objectif d'apprentissage, d'attentes, de cibles, etc. en pensant concomitamment les caractéristiques des tâches et leurs variables didactiques.

9. Il est intéressant de noter que, quasi 20 ans plus tard, dans son ouvrage de 2012, Vial ne parle plus d'évaluation des tâches, mais d'analyse des tâches considérant deux temporalités différentes, à priori (tâche prescrite) et a posteriori (tâche observée et effective).

10. Défi qui reste actuel, plusieurs auteurs considérant que la compétence, par définition, ne peut pas être évaluée.



Rey, dans sa contribution de 2014, discute le rapport controversé entre les compétences et les savoirs. Sans avoir la place de développer ce débat ici, on ne peut cependant pas ne pas l'évoquer dès lors que l'on s'intéresse à l'apprentissage évalué au regard des contextes institutionnels, locaux et internationaux, qui valorisent actuellement les approches par compétences. Savoirs et compétences sont-ils deux réalités incompatibles ou peut-on, comme l'argumente Rey, les considérer comme étant « intriqués les uns dans les autres » (p. 76) ? Sans trancher cette question, on retiendra que, dans l'approche située qui est la nôtre (Mottier Lopez, 2008), que ce soit des compétences ou des connaissances (ou des savoirs si on les place du côté de l'élève), on ne peut pas définir les apprentissages évalués indépendamment des tâches et plus généralement des situations et contextes qui, pour ces derniers, contribuent également à définir les attentes évaluatives. Nous terminons sur ce point.

De l'enjeu des situations et des contextes sociaux pour l'évaluation des apprentissages

Toute production d'élève, tout exercice complété, tout devoir effectué, etc. ne dit rien en soi de sa « valeur ». Dans certaines conceptions de l'évaluation¹¹, l'inférence du « résultat » de l'élève s'associe à un geste de *comparaison* entre l'observable et les objectifs, voire à une mesure d'écart. Pour ma part, je retiens une autre définition de l'évaluation, celle qui insiste sur la fonction de construction de sens et d'attribution de valeur à l'activité de l'élève, depuis les traces qu'il a produites (écrites, orales, motrices) à partir desquelles une interprétation est nécessaire pour appréhender et apprécier l'apprentissage de l'élève, ses acquis, les progressions effectuées, etc. Figari et Remaud (2014) insistent sur le fait que, par le moyen de l'évaluation, un *nouveau* sens est susceptible même d'être produit.

Classiquement, la construction de ce sens est élaborée au regard d'un référentiel formel d'évaluation, sous forme d'objectifs, de critères, parfois d'indicateurs. Ce référentiel est largement lié à la conception (ou modèle s'il existe) du savoir en jeu qui amène, par exemple, à valoriser certaines composantes et critères d'évaluation associés. On pourrait penser, à tort à mon avis, que cette technologie du référentiel, associée à des catégories d'appréciation et échelles (qualitatives, quantitative, de différents ordres) peut suffire pour asseoir l'interprétation de la production de l'élève (toujours au sens large) à des fins de régulation et/ou d'appréciation à formuler. Tâches et activités sont, de fait, toujours saisies dans des « milieux », des « contrats » (au sens didactique des termes) ou dans des contextes sociaux et culturels, dont des « microcultures de classe », elles-mêmes contraintes et rendues possibles par d'autres cultures (Cobb et al., 1997 ; Mottier Lopez, 2008). Ici, le postulat est que l'apprentissage de l'élève ne peut s'appréhender – et donc s'évaluer – que dans la saisie de l'activité de l'élève en interaction avec la tâche, elle-même saisie dans des situations et réseaux

11. Voir par exemple les différentes définitions de l'évaluation dans Figari et Remaud (2014).



de significations plus larges. Le défi de cette conception est considérable, car elle demande d'appréhender ce qui relève du référentiel formel de l'évaluation (objectifs des plans d'études, transposés et reformulés par les enseignant.e.s, les critères d'évaluation définis) et des référents souvent informels associés aux situations, aux contrats, pratiques, normes privilégiées dans les microcultures de classe, en tant qu'arrière-fond interprétatif recelant des significations et des attendus souvent implicites mais qui contribuent également aux processus évaluatifs de construction de sens et de valeur. L'évaluation située se fait alors foncièrement multiréférentielle (Mottier Lopez, 2007). Si nombre de travaux ont donné à voir la variation des jugements évaluatifs des enseignant.e.s et leur subjectivité, peu étudient précisément ces variations en relation avec ces réseaux de significations qui, pourtant, participent pleinement à l'interprétation des productions des élèves sollicitées dans les différentes tâches d'apprentissage et d'évaluation, comme on l'a constaté dans notre étude sur le jugement professionnel de l'enseignant.e (Mottier Lopez & Allal, 2008). Il s'agit à mon sens d'un point important d'élargissement conceptuel de l'évaluation des apprentissages des élèves, ouvrant de nouvelles perspectives de recherche pour toujours mieux comprendre les pratiques évaluatives en contexte.



Références

- Allal, L. & Michel, Y. (1993). Autoévaluation évaluation mutuelle en situation de production écrite. In L. Allal, D. Bain & P. Perrenoud (Ed.), *Evaluation formative et didactique du français* (pp. 239-264). Neuchâtel : Delachaux et Niestlé.
- 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 Publication.
- Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2003). *Assessment for Learning : Putting it into practice*. Oxford, United Kingdom : Oxford University Press.
- Bronckart, J.-P. (1987). Interactions, discours, significations. *Langue française*, 74, 29-50.
- Cobb, P., Gravemeijer, K., Yackel, E., McClain, K. & Whitenack, J. (1997). Mathematizing and symbolizing : The emergence of chains of signification in one first-grade classroom. In D. Kirshner & J. A. Whitson (Ed.), *Situated cognition, social, semiotic, and psychological perspectives* (pp. 151-233). Mahwah : Lawrence Erlbaum Associates.
- Dewey, J. (2011). *La formation des valeurs*. Paris : La Découverte (traduit par A. Bidet, L. Quéré, G. Truc).
- Dolz, J., Schneuwly, B., Thévenaz, Th. & Wirthner, W. (2001). La tâche et ses entours en classe de français. *Actes du congrès*. Neuchâtel : DFLM/IRD (CD-Rom).
- Figari, G. & Remaud, D. (2014). *Méthodologie d'évaluation en éducation et formation*. Bruxelles : De Boeck.
- Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77, 81-112.
- Lave, J. & Wenger, E. (1991). *Situated learning : Legitimate peripheral participation*. Cambridge : Cambridge University Press.
- Laveault, D. & Grégoire, J (2014). *Introduction aux théories des tests en psychologie et en sciences de l'éducation*. Bruxelles : De Boeck.
- Leplat, J. & Hoc, J.-M. (1983). Tâche et activité dans l'analyse psychologique des situations. In J. Leplat (Ed.), *L'analyse du travail en psychologie ergonomique. Tome I* (pp. 47-60). Toulouse : Octarès.
- Margolinas, C. (1993). *De l'importance du vrai et du faux dans la classe de mathématiques*. Grenoble : La Pensée Sauvage.
- Mottier Lopez, L. (2007). L'impact des communautés de pratique dans les entretiens de coévaluation. In A. Jorro (Ed.), *Evaluation et développement professionnel* (pp. 153-167). Paris : L'Harmattan.
- Mottier Lopez, L. (2008). *Apprentissage situé : la microculture de classe en mathématiques*. Berne : Peter Lang.
- Mottier Lopez, L. & Allal, L. (2008). Le jugement professionnel en évaluation : un acte cognitif et une pratique sociale située. *Revue suisse des sciences de l'éducation*, 3, 465-482.
- Piéron, H. (1963). *Examens et docimologie*. Paris : Presses Universitaires de France.
- Rey, B. (2014). *La notion de compétence en éducation et formation*. Bruxelles : De Boeck (Le point sur ... Pédagogie).
- Rey, B., Carette, V., Defrance, A. & Kahn, S. (2003). *Les compétences à l'école : apprentissage et évaluation*. Bruxelles : De Boeck.
- Schneuwly, B. (2001, septembre). Les tâches et leurs entours en classe de français. *Actes du 8^e colloque international de la DFLM*, Neuchâtel.
- Sullivan, P., Clarke, D., Clarke, B. & O'Shea, H. (2010). Exploring the relationship between task, teacher actions, and student learning. *PNA*, 4 (4), 133-142.



- Vial, M. (1996). Classer les tâches. Un complexe entre didactique et évaluation. *Les cahiers de l'année*, 6.
- Vial, M. (2012). *Se repérer dans les modèles de l'évaluation, méthodes dispositifs, outils*. Bruxelles: De Boeck.
- Wiggins, G. P. (1998). *Educative assessment: Designing assessments to inform and improve student performance* [Kindle DX version]. Retrieved from Amazon.com.