

Quelles sont les nouvelles conceptions de l'évaluation en classe ? Vers des évaluations pour apprendre durablement¹



LUCIE MOTTIER LOPEZ

UNIVERSITÉ DE GENÈVE (SUISSE)

Introduction

« L'évaluation est une activité de recueil de diverses informations considérées comme pertinentes par rapport à des objectifs visés, en vue de porter un jugement et de prendre une décision »
(Amigues & Zerbato-Poudou, 1996)

L'étymologie du mot évaluer est « faire ressortir la valeur de », « extraire la valeur de ... ». C'est la fonction pédagogique et institutionnelle qui attribue une connotation sociale à l'évaluation scolaire, notamment au regard du type de décision qu'elle permet de prendre : une décision de régulation, une décision de certification, une décision d'orientation, notamment. Afin d'offrir une ouverture « large » à la problématique de l'évaluation des apprentissages en classe, je vais, dans la première section, rappeler la distinction classique entre évaluation formative et évaluation sommative, et présenter dans les sections suivantes les mises en question de cette distinction et quelques reconfigurations conceptuelles proposées par la littérature scientifique actuelle en sciences de l'éducation et de la formation.

I. À l'origine, une distinction entre évaluation formative et évaluation sommative des apprentissages des élèves

Les travaux internationalement connus de Bloom *et al.* (1971) sur la pédagogie de maîtrise ont amené à différencier l'évaluation *formative* et l'évaluation *sommative* pour ce qui concerne les apprentissages des élèves². Cette distinction a été largement reprise dans tous les continents du monde et elle est, aujourd'hui encore, une référence pour établir une différence entre :

- **Les évaluations dont les résultats sont exploités pour soutenir les apprentissages des élèves.** Dans le *dictionnaire de l'éducation* coordonné par van Zanten, Allal (2008) rappelle que ces évaluations ont une fonction *formative*, c'est-à-dire qu'elles visent « l'adaptation des activités d'enseignement et d'apprentissage afin de favoriser la progression des apprenants vers les objectifs de formation » (p. 312). Les résultats qu'elles produisent fournissent une information à l'enseignant et à l'élève, permettant notamment de différencier les activités didactiques au regard

¹ Pour citer ce document, merci d'utiliser la référence suivante : Mottier Lopez, L. (2023) Quelles sont les nouvelles conceptions de l'évaluation en classe ? Vers des évaluations pour apprendre durablement. Dans Cnesco, *Conférence de consensus du Cnesco l'évaluation en classe, au service de l'apprentissage des élèves : Notes des experts* (pp. 13-22). Cnesco-Cnam.

² Introduite par Scriven, cette distinction existait déjà pour l'évaluation des programmes et des dispositifs.

des besoins des élèves (y compris pour les élèves de bon niveau scolaire) et de repérer des problèmes d'apprentissage et/ou d'enseignement. Elles donnent lieu à des *rétroactions* (ou *feedbacks*) : elles sont exploitées pour réguler (adapter) l'enseignement et soutenir l'autorégulation de l'élève. Elles peuvent être formelles ou informelles, instrumentées (sous forme papier crayon, avec des grilles d'autoévaluation, des listes de vérification, etc.) ou non (dans l'interaction, par l'observation par exemple, mais sans que cela ne donne lieu à des traces consultables par la suite).

- Puis, il y a **les évaluations qui servent à « établir un bilan certifiant les compétences et les connaissances acquises** par l'apprenant à la fin d'une étape intermédiaire (par exemple, semestre) ou d'une étape finale (par exemple, année, cycle pluriannuel) d'un cursus de formation » (Allal, 2008, p. 312). Ces évaluations ont alors une fonction *sommative*. Elles peuvent se dérouler sur toute la durée d'une période de formation (par exemple, par divers « contrôles » ou « tests » ponctuels régulièrement soumis aux élèves donnant lieu à une moyenne numérique ou autre) ou uniquement à la fin (par exemple, un examen final). À noter que les évaluations sont considérées comme sommatives dès que les résultats qu'elles produisent sont pris en compte dans les décisions qui servent à reconnaître institutionnellement les acquis des élèves dans un document officiel, permanent, public³. Elles sont formellement identifiées comme telles : enseignants, élèves, parents des élèves (et tout autre partenaire concerné) savent qu'elles sont associées à des enjeux forts de certification, donc de réussite ou d'échec scolaires, de sélection, d'orientation.

Cette distinction entre évaluation formative et évaluation sommative a fait couler beaucoup d'encre. Certains auteurs considèrent que cette distinction est fonctionnelle car elle permet de reconnaître que, selon leur fonction, les évaluations des apprentissages des élèves ne répondent pas au même contrat scolaire (voir par exemple, Allal, 2008 ; Crahay, Mottier Lopez & Marcoux, 2019 ; Perrenoud, 1998).

En effet, **de façon certes un peu caricaturale**, l'évaluation formative répond à une logique de formation et d'aide à l'apprentissage. Elle *s'intéresse aux erreurs des élèves et à leurs difficultés*, afin de concevoir des dispositifs didactiques et des situations pour les surmonter. Les élèves sont ainsi invités par exemple à expliciter leurs représentations et compréhensions des activités qu'ils réalisent. L'enseignant analyse avec soin leurs réponses, en utilisant notamment des critères d'évaluation (mais pas exclusivement), dans le but de concevoir des activités de remédiation, d'approfondissement, de prolongement. Toujours dans la perspective de mieux accéder aux apprentissages des élèves (y compris des élèves de bon niveau scolaire), ces derniers sont invités à s'impliquer dans l'évaluation formative, pour s'autoévaluer, s'évaluer entre pairs, confronter leur autoévaluation à l'évaluation de l'enseignant, etc. Les dispositifs évaluatifs peuvent être très variés, y compris être intégrés dans les activités d'enseignement et d'apprentissage, c'est-à-dire sans que l'évaluation apparaisse distincte de ces activités. Leur finalité première est de permettre aux élèves de s'autoréguler et de progresser au regard des attendus de l'enseignement et, plus généralement, des programmes de formation et des enjeux pour la formation générale de la personne de l'élève (incluant des compétences sociales, métacognitives, d'organisation, etc.). Autrement dit, l'évaluation formative engage des régulations de l'enseignement et de l'apprentissage au bénéfice de l'élève avant qu'on lui attribue une note à des fins de certification. L'élève a (encore) le droit de se tromper sans conséquence négative pour lui. C'est

³ C'est-à-dire destiné à être légitimement consulté par d'autres personnes que l'enseignant et l'élève concerné.

même parce qu'il fait des erreurs que l'enseignant pourra lui proposer des dispositifs adaptés et que l'apprenant pourra progresser.

L'évaluation sommative répond, quant à elle, à une tout autre logique, surtout quand le système scolaire exige la création de *classements entre les élèves* pour les orienter dans des filières hiérarchisées : pour restreindre l'accès à certaines formations, pour qualifier et disqualifier les élèves en fonction de leurs résultats, etc. Dans un système scolaire *élitiste* qui notamment privilégie des principes associés à une justice méritocratique plutôt qu'à une justice « corrective » visant une égalité des acquis entre élèves par exemple (Crahay, 2012), la logique qui prévaut pour l'évaluation sommative est celle de la sélection et de la compétition : l'évaluation sommative sanctionne les erreurs des élèves en leur attribuant des plus ou moins bonnes notes ; elle établit des classements normatifs qui ont des conséquences tangibles sur la motivation des élèves (positivement ou négativement), sur le climat de compétition entre eux (entre établissements, etc.), sur leur devenir scolaire et plus généralement sur leur formation future et leur insertion dans le monde du travail. Dans ce cas, l'évaluation sommative est « normative » et non pas « critériée ».

Ce sont ces différents éléments qui amènent un ensemble d'auteurs à justifier la distinction à faire entre évaluation formative et évaluation sommative. À leurs yeux, les contrats scolaires sont clairement différents : l'évaluation formative invite les enseignants, les élèves, et tout partenaire impliqué (dont les parents) à un contrat de confiance, de collaboration, de relation d'aide, de « transparence ». Quant à l'évaluation sommative, comme développé par Perrenoud (1998) par exemple, les intérêts entre enseignants et élèves (et leurs parents) ne peuvent que diverger : pour les premiers, il s'agit d'établir de façon la plus précise et correcte possible les niveaux d'acquisition des connaissances et des compétences des élèves, y compris dans la perspective de trier les « méritants » des « moins méritants ». Pour les deuxièmes, il faut « tenter de faire illusion », afin d'être le moins pénalisés possible, se montrer meilleurs que les pairs. La « transparence » et la collaboration ne peuvent être qu'entachées dans l'évaluation sommative par les enjeux sélectifs.

Les parties suivantes de cette note présentent des mises en question de ce raisonnement binaire, entre la « bonne » évaluation (l'évaluation formative) et la « mauvaise » évaluation en raison des pressions et des conséquences « délétères » qu'elle (l'évaluation sommative *normative*⁴, Crahay *et al.*, 2019) génère. Il me paraissait important de rappeler ces premiers développements critiques. Une tendance actuelle en effet est de viser à réconcilier l'évaluation formative et l'évaluation sommative (tendance actuelle à laquelle mes travaux et ceux de mon groupe de recherche contribuent également). Sans quoi, une forme d'angélisme des propositions ci-dessous en sciences de l'éducation pourrait certainement leur être reproché.

Toujours sur un plan théorique, cette note poursuit sur des propositions développées par la littérature scientifique, qui problématisent la distinction évaluation formative / évaluation sommative initialement proposée par Bloom et ses collègues pour les apprentissages des élèves en classe. Deux ensembles de travaux en sciences de l'éducation et de la formation peuvent être identifiés : le premier concerne des travaux menés dans le monde francophone essentiellement et dont on observe une visée de spécification accrue des différentes fonctions de l'évaluation. Le deuxième ensemble, quant à lui,

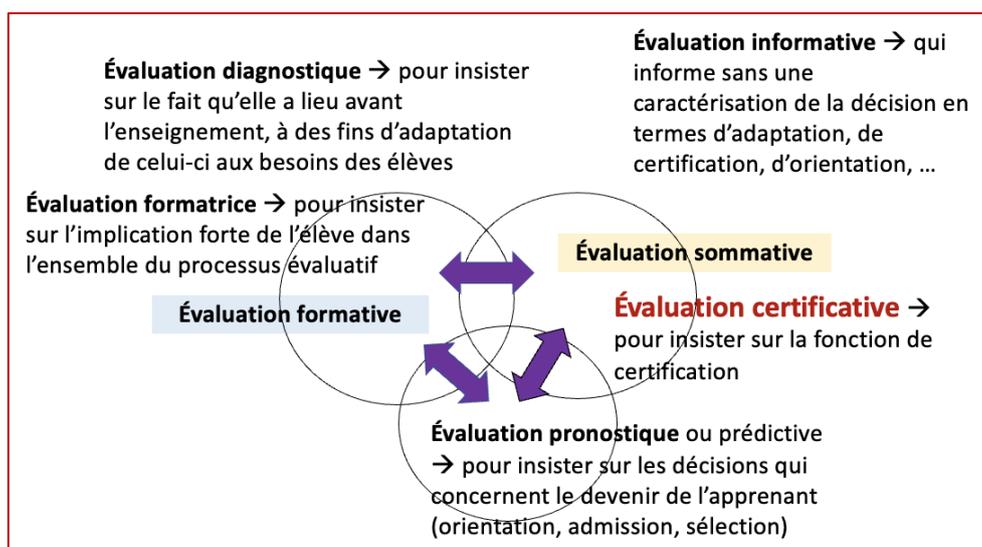
⁴ Cette précision est importante, car il s'agit bien ici d'une évaluation sommative dont les résultats se fondent sur une procédure de comparaison entre les élèves. Comme défini par Crahay *et al.*, (2019), « paraphrasant de Landsheere (1980, p. 295), on définira l'évaluation normative comme suit : toute performance d'élève est évaluée par comparaison avec celles d'individus placés dans la même situation afin de classer le sujet parmi ceux-ci ».

réunit des travaux anglo-saxons principalement qui tendent plutôt à reconceptualiser l'évaluation dans une perspective intégrative visant à dépasser les débats de fonctions. Les sections ci-dessous présentent brièvement ces deux tendances.

II. Vers une spécification toujours accrue des fonctions plurielles de l'évaluation des apprentissages des élèves

Suite aux propositions initiales de Bloom *et al.* (1971), comme je l'ai analysé dans Mottier Lopez (2015a), les auteurs francophones (voir par exemple, Allal, 2008 ; Bonniol & Vial, 1997 ; De Ketele, 2010 ; Hadji, 1989, 2021 ; Vial, 2012), se sont attelés à « décortiquer » les fonctions spécifiques de l'évaluation des apprentissages. Ils s'accordent sur le constat que les seules deux catégories « formative » et « sommative » sont *réductrices* pour rendre compte de la pluralité des buts et des fonctions de l'évaluation. Il est intéressant de les énumérer (Figure 1 ci-dessous) dans la mesure où certaines désignations se retrouvent parfois dans les discours professionnels du terrain et/ou dans des documents institutionnels actuels. Il est à noter toutefois que des variations de définitions existent entre les auteurs (par exemple pour ce qui concerne l'évaluation diagnostique et l'évaluation pronostique entre les propositions de Hadji, 2021, de De Ketele, 2010 ou encore d'Allal, 2008), observées également dans les terrains professionnels et susceptibles d'entraîner des incompréhensions.

Figure 1 : Vers une spécification des différentes fonctions de l'évaluation



La Figure 1, qui est une proposition personnelle, permet de souligner :

- les différentes fonctions (non exhaustives) introduites par la littérature francophone ;
- les proximités entre (1) évaluation formative, évaluation diagnostique, évaluation formatrice, (2) évaluation sommative et évaluation certificative ;
- les relations que les auteurs établissent entre les fonctions (figurées par les flèches et les intersections entre les cercles dans la figure 1).

Comme souligné par Crahay *et al.* (2019), dès les propositions initiales de Bloom et collaborateurs, la relation entre évaluations formative et évaluation sommative a été conceptualisée : l'évaluation formative vise à « préparer » l'évaluation sommative pour permettre la réussite du plus grand nombre

d'élèves dans une perspective de lutte contre l'échec scolaire. Dans la figure 1, les intersections invitent à considérer que les informations / résultats d'une évaluation peuvent potentiellement servir aussi à une autre évaluation, dépassant, ce faisant, les oppositions énoncées plus haut entre elles.

De Ketele (2010) invite, quant à lui, à reconsidérer plus fondamentalement la terminologie attribuée à l'évaluation, critiquant notamment la confusion qu'il considère être faite entre « fonctions » et « démarches ». Le Tableau 1 reproduit la proposition de l'auteur.

Tableau 1 : Croisement entre fonctions et démarches d'évaluation selon De Ketele (2010)

Fonctions → Démarches ↓	Orientation <i>Préparer une nouvelle action</i>	Régulation <i>Améliorer une action en cours</i>	Certification <i>Certifier le résultat d'une action</i>
Sommative <i>Mettre une note</i>	Ex. 1 : Attribuer une note à un test de prérequis langagiers Ex. 2 : Orienter un étudiant sur la base d'un dossier de notes scolaires (comme au terme du collège)	Ex. 1 : Attribuer une note à une interrogation en cours d'apprentissage sur les règles d'accord Ex. 2 : Noter sur 10 points un « quizz test » de 10 questions de compréhension sur un texte en cours d'analyse	Ex. 1 : Attribuer une note à une composition française Ex. 2 : Noter en fin d'année un test composite (analyse, conjugaison, etc.)
Descriptive <i>Identifier et décrire</i>	Ex. 1 : Identifier, parmi 10 prérequis de la lecture, ceux qui sont déjà maîtrisés par un élève du préscolaire Ex. 2 : Identifier dans les questions d'examen de français de l'an dernier les questions qui posent problème	Ex. 1 : Repérer dans la classe les élèves qui n'ont pas réussi un exercice individuel de conjugaison pour organiser un tutorat Ex. 2 : Organiser un test diagnostique sur un apprentissage important	Ex. 1 : Remplir un bulletin descriptif (cocher dans une liste d'objectifs du cours de français ceux qui sont atteints de façon stable) Ex. 2 : Rédiger un certificat mentionnant et décrivant les compétences langagières acquises
Herméneutique <i>Donner un sens</i>	Ex. 1 : Orienter un étudiant vers une filière littéraire sur la base de nombreux indices (résultats scolaires, observations, entretien) Ex. 2 : Analyser finement en début d'année une première dictée et un premier texte libre des nouveaux élèves en vue de préparer les premières séquences	Ex. 1 : Observer et analyser les productions des élèves pendant une séquence de travail individuel en classe Ex. 2 : Analyser les résultats d'un test diagnostique à la lumière d'une analyse de la façon dont les séquences d'apprentissage antérieures ont été réalisées	Ex. 1 : Préparer une réunion de parents sur la base d'un portfolio des productions significatives de ses élèves Ex. 2 : Certifier le passage de classe, non seulement sur la base des notes obtenues en fin d'année, mais en tenant compte d'autres indices pertinents

De ce tableau, on retiendra notamment que c'est la désignation « évaluation certificative » qui est privilégiée plutôt qu'évaluation sommative. D'une part, elle désigne plus précisément la fonction de cette évaluation (certifier institutionnellement les acquis des élèves), alors que, d'autre part, le terme « sommatif » est associé à l'idée de faire un bilan en tant que démarche pouvant servir les fonctions d'orientation, de régulation, de certification. Quant à l'évaluation formative, elle apparaît liée à la

fonction de régulation. La fonction d'orientation est quant à elle associée tant à l'évaluation diagnostique qu'à l'évaluation pronostique (voir Figure 1 pour un rappel de celles-ci).

On retiendra de ce premier ensemble de travaux que les auteurs ont argumenté en faveur de terminologies plus spécifiques et pointues, au point peut-être de risquer un certain *éclatement* entre les différentes fonctions, types, démarches évaluatives. La proposition de De Ketele montre sa préoccupation à penser les liens entre les éléments en faveur d'une approche plus intégrative et systémique, dépassant largement la dichotomie « formatif – sommatif » initiale.

Sur le plan conceptuel, ce travail de spécification théorique rend compte (en partie) de la complexité des pratiques évaluatives et des enjeux scientifiques pour tenter de toujours mieux les comprendre et les caractériser. Des résultats de recherches tant anglophones que francophones nous alertent cependant : ces distinctions conceptuelles restent peu opérationnelles dans les pratiques en classe (voir par exemple, Abernot, 2013 ; Black & Wiliam, 1998 ; Mottier Lopez, 2015b ; Perrenoud, 1991). Par exemple, au regard des cas de figure suivants (non exhaustifs) :

- Les enseignants tendent à ne pas faire la différence entre les fonctions de l'évaluation : par exemple, une évaluation est vue comme pouvant servir à « mettre une note » et à la fois à être utilisée pour envisager des activités de remédiation, pour planifier une prochaine séquence d'enseignement dans un but formatif *a posteriori*. Une même évaluation sert plusieurs fonctions : « pourquoi alors les différencier ? »⁵. Cette indifférenciation dans la pratique prend le dessus sur les conceptualisations théoriques qui font peu sens face à la réalité des mises en œuvre *in situ*.
- L'évaluation reste principalement perçue et représentée au regard de ses enjeux forts de certification, voire d'orientation et de sélection. L'évaluation formative est peu ou n'est pas pratiquée ou, quand elle est interactive et intégrée dans les situations d'enseignement et d'apprentissage (voir par exemple, Allal & Mottier Lopez, 2005 ; Mottier Lopez, 2015b), elle n'est pas vue comme étant une évaluation par les enseignants : « c'est pas évaluer, c'est enseigner ! ». La représentation de l'évaluation ici reste celle d'une pratique formelle, ponctuelle, instrumentée, débouchant nécessairement sur une « note » (qualitative ou quantitative). Tout ce qui s'en écarte n'est pas conçu comme une évaluation. La conséquence est que l'évaluation formative dite « élargie » (intégrant des modalités interactives informelles, impliquant l'élève dans des démarches d'autoévaluation au sens large, etc. voir Allal & Mottier Lopez, 2005 pour une description étayée) n'est pas reconnue en tant que telle.
- Les enseignants considèrent que les fonctions formatives et sommatives / certificatives sont inconciliables ; « parfois on doit être coach, parfois on doit être juge ! » et ils ne le vivent pas bien. Faire cohabiter des fonctions différentes de l'évaluation dans la classe exige de parvenir à gérer des tensions complexes, parfois paradoxales, notamment pour ce qui concerne les enjeux de certification et d'orientation (Cardinet, 1984 ; Harlen, 2007 ; Laveault, 2008).

Est-ce alors une bonne « idée » de différencier de façon aussi pointue les différentes fonctions et démarches évaluatives, ou serait-il préférable de concevoir des approches plus intégratives, caractérisée par la *finalité éducative visée* plus que par les différentes décisions qui découlent de

⁵ J'insère ici des exemples de phrases souvent entendues dans mes interventions de formation ou de recherches collaboratives. À noter dans le cas de figure rapporté ici que l'aide à l'apprentissage de l'élève se fait alors après la note...

l'évaluation (régulation, certification, orientation) ? C'est l'option tendancielle retenue par les travaux cités ci-dessous.

III. Vers des propositions de reconfiguration conceptuelle visant une « évaluation pour apprendre durablement »

Depuis les années 1990, un ensemble de travaux essentiellement anglophones ont tenté de ne plus opposer les différentes fonctions de l'évaluation, en faveur d'approches plus intégratives et systémiques. De nouvelles « modélisations » ont été proposées, synthétisées dans le Tableau 2 ci-dessous.

Tableau 2 : Propositions conceptuelles visant à reconfigurer l'évaluation des apprentissages (Mottier Lopez, Girardet & Naji, 2021, p. 6, augmenté)⁶

Terme en anglais	Auteurs à l'origine	Traduction en français	Dimensions principales mises en évidence
<i>Assessment for learning</i>	<i>Assessment Reform Group</i> (Black & Wiliam, 1998)	Évaluation pour apprendre Évaluation soutien d'apprentissage (Allal & Laveault, 2009)	Englobe toutes formes d'évaluations (formelles et informelles, formatives et sommatives) susceptibles d'encourager les élèves à s'engager dans des apprentissages à des fins de régulation et d'autorégulation ; fortement intégrées aux activités d'enseignement et d'apprentissage.
<i>Authentic assessment</i>	Wiggins (1998)	Évaluation authentique	Les contenus, la structure, les caractéristiques et la validité des évaluations s'associent à des situations complexes du monde « réel », sans distinction entre évaluations formatives et certificatives.
<i>Sustainable assessment</i>	Boud (2000)	Évaluation durable	Les évaluations, quelles que soient leurs formes et leurs fonctions, répondent aux besoins d'apprentissage dans des temporalités courtes et longues (apprentissage tout au long de la vie).
<i>Assessment as learning</i>	Earl (2003)	Évaluation comme apprentissage	Les évaluations sont vues comme susceptibles de représenter des occasions d'apprentissage, particulièrement quand les élèves sont impliqués dans des autoévaluations, des évaluations entre pairs et des évaluations collaboratives formatives. L'évaluation devient objet d'apprentissage.
<i>Learning-oriented assessment</i>	Carless, Joughin et Mok (2006)	Évaluation axée sur l'apprentissage	Des évaluations qui cherchent à réconcilier les fonctions formatives et sommatives pour que toutes orientent l'apprentissage des élèves dans des directions productives.

Un point commun à toutes ces propositions conceptuelles, tant pour les évaluations formatives que sommatives/certificatives, est de les concevoir toutes comme des leviers potentiellement *positifs* pour soutenir des apprentissages de qualité chez les élèves. Quelle que soit sa fonction, l'évaluation se veut « constructive » et « soutenante » pour les apprentissages des élèves et leur motivation à apprendre, sur des temporalités ponctuelles et plus longues. Comme écrit dans Mottier *et al.* (2021), à quelques

⁶ Les références citées dans le tableau sont celles, notamment, à l'origine des modèles exposés. D'autres contributions ont ensuite largement développé les propositions y compris dans la littérature scientifique actuelle. Je ne les cite pas ici. Elles sont très nombreuses.

nuances près, les auteurs francophones et anglophones dans la littérature *actuelle*⁷ s'accordent alors sur un ensemble de caractéristiques à problématiser. Notamment :

- La cohérence entre ce qui est enseigné, appris et évalué pour construire un « alignement curriculaire » cohérent, y compris avec les didactiques disciplinaires et les prescriptions institutionnelles ;
- La nature des tâches d'apprentissage qui sont susceptibles d'intégrer des évaluations associées à des possibilités de régulation et d'autorégulation des élèves, qui encouragent des collaborations entre eux et avec l'enseignant ;
- Les contrats évaluatifs en jeu, les attentes, les objectifs, les critères d'évaluation, y compris pour renforcer la confiance mutuelle entre apprenants et avec les enseignants au regard des différents enjeux de l'évaluation ;
- L'implication active des élèves dans des démarches d'autoévaluation et d'évaluation entre pairs, y compris dans la perspective de développer chez les élèves des compétences évaluatives, de jugement critique et d'autorégulation ;
- La façon de donner de l'importance et un rôle à la « voix des élèves » pour les considérer comme des partenaires à part entière dans l'évaluation et ses enjeux ;
- Les gestes professionnels de l'enseignant, ses postures évaluatives, son jugement professionnel et les modérations sociales entre professionnels pour construire des interprétations et des décisions évaluatives les plus pertinentes et valides possibles ;
- Les opportunités de rétroaction (feedback) de la part des enseignants et entre élèves dans la perspective de soutenir les apprentissages dans les tâches concernées, incluant les processus de régulation, de co-régulation, d'autorégulation ;
- Les cultures de classe et d'établissement scolaires, construites avec les élèves et l'ensemble des partenaires concernés, en faveur « d'évaluations positives pour l'avenir » (*sustainable assessment*), c'est-à-dire dans la perspective d'une évaluation « pour apprendre tout au long de la vie ».

La force des propositions issues de cette littérature est notamment qu'elles se fondent sur des données empiriques produites avec des protocoles de recherche expérimentaux, des observations, des enquêtes, des analyses comparatives, des interventions.

⁷ Très nombreux, je ne suis pas en mesure de les citer dans cette note brève. Je propose par exemple de consulter l'ouvrage collectif de Barroso da Costa, Nizet et Leduc (2021) qui propose des synthèses étayées de quarante ans de travaux en évaluation.

Références

- Allal, L. (2008). Évaluation des apprentissages. In A. van Zanten (Ed.), *Dictionnaire de l'éducation* (pp. 311-314). Presses Universitaires de France.
- Allal, L. & Laveault, D. (2009). Assessment for learning: évaluation-soutien d'apprentissage. *Mesure et évaluation en éducation*, 32(2), 99-106. <https://doi.org/10.7202/1024956ar>
- Allal, L. & Mottier Lopez, L. (2005). L'évaluation formative de l'apprentissage : revue de publications en langue française. Dans OCDE, *L'évaluation formative : pour un meilleur apprentissage dans les classes secondaires* (p. 265-290). Éditions OCDE/CERI.
- Amigues, R. & Zerbato-Poudou, M.-T. (1996). *Les pratiques scolaires d'apprentissage et d'évaluation*. Dunod.
- Barroso da Costa, C., Nizet, I. & Leduc, D. (Eds.) (2021). *Quarante ans de mesure et d'évaluation*. Presses Universitaires du Québec.
- Black, P. & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148. <http://electronicportfolios.org/afl/InsideBlackBox.pdf>
- Bloom, B. S., Hasting, J. T. & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. McGraw-Hill.
- Bonniol, J.-J. & Vial, M. (1997). *Les modèles de l'évaluation*. De Boeck.
- Boud, D. (2000). Sustainable assessment: rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151-167. doi: 10.1080/71369572
- Cardinet, J. (1984). *Pour apprécier le travail des élèves*. Institut romand de recherches et de documentation pédagogiques.
- Crahay, M. (2012). *L'école peut-elle être juste et efficace ?* De Boeck.
- Crahay, M., Mottier Lopez, L. & Marcoux, G. (2019). L'évaluation des élèves : Docteur Jekyll and Mister Hyde de l'enseignement. In M. Crahay (Ed.), *Peut-on lutter contre l'échec scolaire ?* (pp. 358-425). De Boeck.
- De Ketele, J. (2010). Ne pas se tromper d'évaluation. *Revue française de linguistique appliquée*, XV (1), 25-37. doi:10.3917/rfla.151.0025.
- Earl, L. M. (2003). *Assessment as Learning. Using assessment to maximize student learning*. Hawker Brownlow Education.
- Hadji, C. (1989). *L'évaluation, règles du jeu*. ESF éditions.
- Hadji, C. (2012). *Faut-il avoir peur de l'évaluation ?* De Boeck.
- Hadji, C. (2012). *Le défi d'une évaluation à visage humain*. ESF.
- Harlen, W. (2007). *Assessment of learning*. SAGE publications ltd.
- Laveault, D. (2008). Le jugement professionnel : foyer de tensions et de synergies nouvelles en évaluation scolaire. *Revue suisse des sciences de l'éducation*, 30 (3), 483-500.
- Mottier Lopez, L. (2015a). *Évaluations formative et certificative : enjeux pour l'enseignement*. De Boeck.

Mottier Lopez, L. (2015b). Évaluation-régulation interactive : étude des structures de participation guidée entre enseignant et élèves dans le problème mathématique 'Enclos de la chèvre'. *Mesure et évaluation en éducation*, 38(1), 89-120.

Mottier Lopez, L. (2015c). L'évaluation formative des apprentissages des élèves : entre innovations, échecs et possibles renouvellements par des recherches participatives. *Questions Vives*, 23, 31-50.

Mottier Lopez, L., Girardet, C. & Naji, T. (2021). L'évaluation continue pour apprendre : enjeux de la pluralité des feedbacks entre pairs dans un cours universitaire. *Mesure et évaluation en éducation*, 44(2), 1-34.

Perrenoud, P. (1991). Pour une approche pragmatique de l'évaluation formative. *Mesure et évaluation en éducation*, 13(4), 49-81.

Perrenoud, P. (1998). *L'évaluation des élèves. De la fabrication de l'excellence à la régulation des apprentissages*. De Boeck.

Vial, M. (2012). *Se repérer dans les modèles de l'évaluation*. De Boeck.

Wiggins, G. P. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. Jossey-Bass.