UNIVERSITÉ DE GENÈVE

Archive ouverte UNIGE
https://archive-ouverte.unige.ch

Article scientifique | Article | 2015

Published version | Open Access

This is the published version of the publication, made available in accordance with the publisher's policy.

L'évaluation formative des apprentissages des élèves : entre innovations, échecs et possibles renouveaux par des recherches participatives

Mottier Lopez, Lucie

How to cite

MOTTIER LOPEZ, Lucie. L'évaluation formative des apprentissages des élèves : entre innovations, échecs et possibles renouveaux par des recherches participatives. In: Questions vives, 2015, n° 23. doi: 10.4000/questionsvives.1692

This publication URL: https://archive-ouverte.unige.ch//unige:80948
Publication DOI: 10.4000/questionsvives.1692

© This document is protected by copyright. Please refer to copyright holder(s) for terms of use.

Questions vives

Questions Vives

N° 23 (2015)
Évaluation et changement

Lucie Mottier Lopez

L'évaluation formative des apprentissages des élèves : entre innovations, échecs et possibles renouveaux par des recherches participatives

Avertissement
Le contenu de ce site relève de la législation française sur la propriété intellectuelle et est la propriété exclusive de l'éditeur.
Les œuvres figurant sur ce site peuvent être consultées et reproduites sur un support papier ou numérique sous réserve qu'elles soient strictement réservées à un usage soit personnel, soit scientifique ou pédagogique excluant toute exploitation commerciale. La reproduction devra obligatoirement mentionner l'éditeur, le nom de la revue, l'auteur et la référence du document.
Toute autre reproduction est interdite sauf accord préalable de l'éditeur, en dehors des cas prévus par la législation en vigueur en France.

revues.org

Revues.org est un portail de revues en sciences humaines et sociales développé par le Cléo, Centre pour l'édition électronique ouverte (CNRS, EHESS, UP, UAPV).

Référence électronique
Lucie Mottier Lopez, « L'évaluation formative des apprentissages des élèves : entre innovations, échecs et possibles renouveaux par des recherches participatives », *Questions Vives* [En ligne], N° 23 | 2015, mis en ligne le , consulté le 14 février 2016. URL : http://questionsvives.revues.org/1692 ; DOI : 10.4000/questionsvives.1692

Éditeur : Université de Provence - Département des Sciences de l'éducation
http://questionsvives.revues.org
http://www.revues.org

Document accessible en ligne sur :
http://questionsvives.revues.org/1692
Document généré automatiquement le 14 février 2016. La pagination ne correspond pas à la pagination de l'édition papier.
Tous droits réservés

Lucie Mottier Lopez

L'évaluation formative des apprentissages des élèves : entre innovations, échecs et possibles renouveaux par des recherches participatives

1 Dans les années 1970-2000, l'évaluation formative des apprentissages des élèves a suscité un formidable élan d'intérêt, tant au sein des communautés scientifiques que dans les politiques éducatives en Suisse romande. Pourtant, aujourd'hui, cela ne semble plus aller autant de soi de parler d'évaluation formative ou, alors, sous d'autres vocables. Comment comprendre cet intérêt puis désintérêt ? À partir de synthèses de la littérature et de bilans produits par des instances, les deux premières parties de l'article retracent les grandes lignes du contexte socio-historique à propos de l'évaluation formative des apprentissages des élèves en Suisse romande : d'abord, l'intérêt pour cette fonction de l'évaluation au plan scientifique et dans le contexte politique qui a débouché sur des décisions de réformes scolaires ; puis, l'échec des changements souhaités avec un retour en force de l'évaluation certificative chiffrée dans les écoles primaires.

2 Face à ces constats, doit-on baisser les bras et accepter que l'évaluation scolaire soit toujours plus sélective et marquée par la culture ambiante de la mesure, du chiffre, du « rendre compte » ? Ou, à partir des expériences passées, des possibilités restent-elles envisageables pour construire des espaces de changement en faveur de pratiques évaluatives qui se mettent au service des apprentissages des élèves et de la différenciation pédagogique ? La troisième partie de l'article réfléchit à des possibilités de transformation des pratiques évaluatives des enseignants par le moyen de recherches participatives et collaboratives. Toutes inscrites dans le contexte de l'école primaire genevoise, trois recherches sont présentées. Leurs visées sont discutées au regard du contexte socio-historique dans lequel elles se situent et par rapport à des enjeux de développement professionnel des enseignants. Dans cet article, le but n'est pas de présenter en soi les résultats produits par ces recherches. Mais, par rapport au thème de ce numéro de revue, l'idée est d'objectiver la forme d'engagement des enseignants dans ces recherches vis-à-vis de la question sensible de l'évaluation des apprentissages en classe à des fins transformatives.

1. La montée en puissance de l'évaluation formative en Suisse romande

3 Cette partie dégage les principaux mouvements qui ont contribué à promouvoir l'évaluation formative des apprentissages des élèves en Suisse romande et qui ont débouché sur des décisions politiques de réformes scolaires. Elle se fonde sur des analyses systématiques que nous avons effectuées de la littérature de recherche de langue française sur l'évaluation formative entre 1979 et 2002 dans Allal et Mottier Lopez (2005), complétée jusqu'en 2007 et mise en perspective avec des enjeux d'évaluation certificative dans Mottier Lopez et Laveault (2008), puis complétée de 2008 à 2014 dans une communication visant à réfléchir les liens conceptuels entre évaluation, apprentissage et didactique dans Mottier Lopez (2015, janvier). Ces synthèses ont été croisées avec des écrits institutionnels produits en Suisse romande dans la période des réformes scolaires entre 1980 et 2010. Dans une logique d'analyse documentaire, et afin d'illustrer les grandes orientations thématiques, nous citons dans cette partie les titres d'ouvrages ou de rapports qui ont particulièrement marqué les contextes scientifiques et politiques qui nous intéressent.

1.1. L'évaluation formative, une question vive pour la communauté scientifique

4 En Suisse romande, une « problématique évaluative » à propos des apprentissages des élèves a commencé à voir le jour dans les années 1970, suite notamment aux propositions de la « pédagogie de maîtrise » (*mastery learning*, Bloom, Hastings & Madaus, 1971) diffusées et débattues par des chercheurs américains en poste à l'Université de Genève, Linda Allal et Michael Huberman. La distinction entre évaluation sommative et formative, développée dans les travaux de Bloom pour l'évaluation des apprentissages des élèves[1], amène Cardinet (1983), alors mandaté pour construire des épreuves d'orientation, à argumenter une troisième fonction : l'évaluation prédictive, ensuite nommée évaluation pronostique. Celle-ci vise à distinguer les pratiques évaluatives qui servent plus spécifiquement à l'orientation et à l'admission des élèves dans des programmes de formation. Perrenoud (2001) définit comme suit ces différentes fonctions[2] :

L'évaluation formative. Sa fonction est de soutenir la régulation des enseignements et des apprentissages en train de se faire ; elle se déploie à l'intérieur d'un cursus scolaire ;

L'évaluation certificative. Elle garantit des acquis à l'égard de tiers, sur le marché du travail, à la rigueur à la fin d'un cycle d'études ; elle intervient à l'issue d'un cursus donné ;

L'évaluation pronostique. Elle fonde des décisions de sélection ou d'orientation en fonction de l'aptitude présumée à suivre un nouveau cursus, par exemple telle filière du secondaire ; elle se situe en amont d'un cursus et sous-tend un choix. (p. 19)

5 À l'instar d'autres pays francophones, le débat sur les différentes fonctions de l'évaluation est vif, notamment au sujet de l'idée innovante, à l'époque, que l'évaluation pouvait se mettre au service de l'apprentissage de l'élève et de l'amélioration des dispositifs d'enseignement. L'ouvrage collectif, *L'évaluation formative dans un enseignement différencié*, publié par Allal, Cardinet et Perrenoud en 1979, rend compte de cet intérêt pour l'évaluation formative en Suisse romande. Il émane d'échanges entre chercheurs belges et suisses entre 1977 et 1978[3], suite au constat d'un manque de clarté à propos des concepts utilisés, d'une confusion entre terminologies parfois divergentes et de résultats mal connus à propos d'expériences souvent isolées. Des échanges internationaux s'organisent entre la Belgique, la France et la Suisse, débouchant sur la création, en 1986, de l'Association francophone pour le développement des méthodologies d'évaluation en éducation en Europe (ADMEE-Europe). À l'occasion du premier colloque à Dijon, une controverse marque les travaux, reprise au colloque de 1987 : dans quelle mesure l'évaluation formative doit-elle être plutôt investie par la recherche en didactique des disciplines ou peut-elle faire l'objet d'un domaine de recherche à part entière ?

6 Cette question, en Suisse romande, restera vive dans les années 1980 et 1990, psychologues, didacticiens et sociologues partageant un intérêt commun pour cette nouvelle forme d'évaluation des apprentissages des élèves. L'ouvrage collectif d'Allal, Bain et Perrenoud, *Évaluation formative et didactique du français* (1993), en témoigne. Il ose l'idée que l'évaluation formative, qui demande à prendre en considération à la fois les contenus de savoir enseignés et les processus psychologiques d'apprentissage des élèves, pourrait mettre fin à la séparation classique entre didactique et évaluation[4]. Quant au regard sociologique de Perrenoud, il examine l'évaluation formative et la différenciation pédagogique, son corollaire, comme des stratégies susceptibles de lutter contre l'échec scolaire et, plus largement, de contribuer à une politique de démocratisation des études.

7 Plus généralement, cet intérêt pour l'évaluation formative débouche, dans la communauté scientifique francophone, sur de nombreux travaux qui contribuent à un élargissement significatif de son champ conceptuel (Allal & Mottier Lopez, 2005). Une conception élargie[5] de l'évaluation formative se développe progressivement par rapport au modèle de la pédagogie de maîtrise, prenant de la distance, voire créant plus radicalement une rupture épistémologique, avec son ancrage post-béhavioriste initial.

1.2. D'un état des lieux des pratiques aux décisions de réformes en Suisse romande

8 Concomitamment aux travaux scientifiques, en Suisse romande, la commission pédagogique de la conférence suisse des directeurs cantonaux de l'instruction publique entreprend un examen de la situation de l'école primaire entre 1980 et 1984 (projet SIPRI). Ce projet s'appuie sur des groupes de travail et un réseau d'écoles de contact pour élaborer et expérimenter des « améliorations » dans différents domaines, dont celui de « l'appréciation du travail des élèves ». Le rapport conclusif de Salamin (1986) insiste sur l'importance pour l'enseignant de distinguer trois fonctions à l'évaluation (formative, sommative, pronostique). Il souligne la valeur des démarches d'autoévaluation menées par les élèves et l'importance de la communication aux parents. Il recommande également d'espacer les temps de promotion et de promouvoir activement l'évaluation formative dans les classes.

9 Ce rapport aura un impact sur l'orientation des politiques éducatives des cantons romands et des réformes qui suivront. Dans les années 1990, une commission romande est mandatée par les responsables politiques des cantons romands pour formuler des propositions visant à faire progresser les approches et pratiques de l'évaluation dans les divers degrés de la scolarité obligatoire. Un rapport est rédigé par Weiss (1996), *Vers une conception cohérente de l'évaluation pour la scolarité obligatoire en Suisse romande et au Tessin*. S'en suit la tenue en 1997 d'un colloque romand sur l'évaluation, organisé sous l'égide de la conférence intercantonale de l'instruction publique de la Suisse romande et du Tessin (seul et unique colloque sur ce thème qui a eu lieu jusqu'à aujourd'hui). Il débouche sur des actes au titre révélateur : *Pour une évaluation plus formative* (Weiss, Wirthner, Bain, Greub & Perrenoud, 1998).

10 D'une façon générale, comme l'observe le bilan rétrospectif de Wirthner (2010), les années 1980-1990 servent à constituer un ensemble de connaissances sur l'évaluation formative. Les assises théoriques et également politiques, fondées sur des premières expériences dans des écoles pilote, semblent alors réunies pour entreprendre des réformes en matière d'évaluation scolaire et plus largement. Le canton de Genève ouvre la marche, avec une phase « d'exploration intensive » de la réforme, appelée *Rénovation*, de l'école primaire entre 1994-1998, puis une phase « d'extension progressive » entre 1998-2002. Dans le canton de Vaud, à la suite d'une votation référendaire, la réforme *École vaudoise en mutation* est acceptée en 1997, en réaction à un système d'évaluation et d'orientation jugé trop sélectif. Bien qu'étant importantes, les transformations de l'évaluation scolaire ne représentent, dans ces deux réformes, qu'un aspect parmi d'autres changements majeurs, dont une réorganisation de l'école en cycles d'apprentissage pluriannuels. Le changement des pratiques et le changement des structures sont jugés comme étant co-dépendants.

2. Plébiscite pour un retour à une évaluation certificative chiffrée

11 L'ouvrage collectif de Gilliéron Giroud et Ntamakiliro (2010) établit un bilan des réformes dans le domaine de l'évaluation en Suisse romande. Le bilan est plutôt mitigé, pour ne pas parler d'échec. Cette partie de l'article expose les grandes lignes de résistance puis esquisse quelques pistes interprétatives.

2.1. Confusion entre évaluations formative et certificative, et difficultés techniques

12 Dès le milieu des années 2000, c'est-à-dire à peine après quelques années de mise en œuvre des réformes, un climat social conflictuel se développe à propos essentiellement de la suppression des notes chiffrées à l'école primaire. Comme l'observe Blanchet (2010), en référence aux travaux d'Allal et al. (1979) :

> L'abandon des notes devait permettre de supprimer un stress handicapant pour de nombreux élèves, de centrer les élèves et les maîtres sur les contenus à acquérir et de libérer un temps important et précieux pour de multiples régulations favorisant l'apprentissage plutôt que la sélection. (pp. 83-84)

13 Mais ces visées sont mal comprises. L'accent mis sur l'évaluation formative laisse croire que la certification des acquis des élèves en fin de cycles d'apprentissage n'existe plus. Un hiatus se développe auprès des partenaires sociaux (y compris chez les enseignants) qui ne comprennent pas la fonction certificative d'un livret scolaire sans notes dans les premiers degrés de l'école primaire. Autrement dit, l'évaluation certificative qui se transmet de façon qualitative en fin de cycle est considérée comme une évaluation formative. L'analyse de Blanchet (2010), dans le canton de Vaud, expose aussi les difficultés liées à l'usage d'appréciations qualitatives dans une évaluation certificative : lourdeur des procédures, choix des termes pour remplacer les notes, construction d'un sens aux appréciations qualitatives, appréhension de l'apprentissage et de la réussite par rapport aux exigences de promotion, retour à des échelles, combinaisons d'appréciations qualitatives et quantitatives, établissement d'un seuil de réussite sans calculer une moyenne. Blanchet (2010) constate un manque d'anticipation à propos des aspects plus techniques et pratiques pour savoir comment évaluer les apprentissages sans utiliser des notes et des moyennes.

14 Globalement, les procédures proposées aux enseignants étaient lourdes à mettre en place ; les parents ont perdu leurs repères, considérant paradoxalement que la suppression des notes et moyennes chiffrées renforçait la sélectivité de l'école (Wirthner, 2010). On voit ici combien les deux « missions » de l'évaluation, au sens d'Abernot (2013), sont difficiles à démêler, entre une évaluation qui se met au service de la formation de la personne et une évaluation qui a une fonction de tri social. Constatant ces résistances, et les débats passionnés qui en découlent, les cantons de Fribourg et du Valais renoncent à réformer leur système d'évaluation. Dans le canton de Genève, un vote populaire en 2006 plébiscite la réintroduction des notes au deuxième cycle de l'école primaire (élèves de 8 à 12 ans) et une décision de promotion à la fin de chaque degré scolaire. Le conseiller d'État de l'époque, en charge du département de l'instruction publique, durcit le vote en ré-instituant le calcul de moyennes de notes pour établir la certification annuelle. Dans le canton de Vaud, une initiative populaire sollicite également l'usage des notes dès la troisième année primaire[6] (élèves de 8 ans). Un contre-projet est accepté en 2005, après un cycle de forums et de débats sociaux, qui repousse les notes chiffrées à la cinquième année primaire (élèves de 10 ans).

15 L'évaluation formative reste identifiée dans les cadres réglementaires des différents cantons romands. Mais on constate aujourd'hui qu'elle ne représente plus réellement une priorité ni dans les discours politiques, ni dans les offres de formation continue professionnelle du canton de Genève comme notre analyse documentaire l'a mis en évidence dans Mottier Lopez (2014). Sept ans après le retour des notes à l'école primaire genevoise, un renforcement s'observe même. En effet, l'usage de demi-notes et le calcul de moyennes au dixième d'un ensemble de notes a été décidé en automne 2013. Quant aux deux premiers degrés de la scolarité (élèves de 4-6 ans), les enseignants doivent désormais formuler des appréciations par domaine disciplinaire dès la fin du 2ème degré (élèves de 6 ans). Auparavant, seule une appréciation globale de la progression de l'élève était demandée à cette étape de la scolarité.

2.2 Un mouvement (relatif) de désintérêt face à l'évaluation formative

16 Dans les discours politiques actuels en Suisse romande, force est de constater que l'évaluation formative ne représente plus (autant) un enjeu de transformation des pratiques enseignantes. La confusion entre évaluation formative et évaluation certificative sans notes a quasiment discrédité la fonction pédagogique et didactique de l'évaluation consistant à fournir des informations pour soutenir l'apprentissage des élèves et pour réguler et différencier les processus d'enseignement. Radicalement dit, dès qu'une évaluation sans notes est évoquée (ce qui est le cas de l'évaluation formative), il y a suspicion de subjectivité arbitraire et d'injustice. La symbolique véhiculée par les notes (et moyennes de notes) reste profondément ancrée dans les représentations sociales : le chiffre serait garant de plus de justice, d'objectivité, de fiabilité. Certains pensent même que la culture du chiffre, des indicateurs chiffrés, de la mesure, du rendre compte (*accountability*) envahit toujours plus la sphère éducative, marquant de son empreinte les conceptions évaluatives (Felouzis & Hanhart, 2011). L'évaluation formative

n'est plus un sujet politique porteur dans une société qui tend à renforcer les processus de sélection, bien que laissant croire au mythe de l'égalité des chances.

17 Ce recul de l'évaluation formative ne s'observe pas seulement sur un plan politique. En effet, dans la communauté scientifique francophone également, notre analyse de la littérature dans Mottier Lopez (2015, janvier) montre que l'évaluation formative des apprentissages des élèves ne représente plus un domaine d'investigation aussi vif que dans les années 1980-2000.[7] Le terme *évaluation formative* est parfois perçu comme galvaudé, y compris dans la communauté anglophone (Wiliam, 2011). On parlera alors d'évaluation soutien d'apprentissage, d'évaluation authentique, d'évaluation alternative, etc. Plus radical, Perrenoud (1998) se demandait déjà à la fin des années 1990 s'il ne fallait pas éviter le terme d'*évaluation* quand le projet est de l'associer à l'apprentissage, la régulation, l'autorégulation… tant celui-ci reste connoté dans les représentations sociales à la mesure et au contrôle normatif.

18 Avons-nous fait fausse route en envisageant une fonction formative de développement à l'évaluation des apprentissages qui devrait être prioritaire face à la mission de sélection qui est également attribuée à l'école ? Comme y croit Abernot (2013), peut-on continuer de plaider pour une évaluation qui :

> est à la fois un moyen et une fin. Un moyen, car pour avancer, il faut constamment être renseigné de ses progrès et se sentir valorisé par eux. Mais c'est aussi une fin car c'est l'évaluation qui donne à l'homme son potentiel d'évolution. (p. 196)

19 Dans cette perspective, quelles pistes entrevoir pour continuer de penser le changement des conceptions, des pratiques, des théorisations d'une évaluation au service avant tout de la formation des personnes, sans pour autant nier les besoins de validation et de reconnaissance sociales de leurs acquis ?

3. Construire avec les évènements passés pour penser le changement des pratiques évaluatives de demain

20 Transformer les pratiques enseignantes et, plus largement, changer un tant soit peu la culture scolaire et le rapport à l'école est un défi. Comme on l'a vu, l'évaluation scolaire représente un de ces objets cruciaux autour desquels s'expriment les tensions face aux différentes missions de l'école et de son rôle dans l'éducation et la société (Blanchet, 2010). Les discours idéologiques à propos de l'évaluation sont forts, tant en sa faveur qu'en sa défaveur. Dans Allal et Mottier Lopez (2005), l'analyse des travaux francophones sur l'évaluation formative entre 1979 et 2002 a montré que l'évaluation formative s'est développée essentiellement sur un plan théorique. Le constat était que « les promesses théoriques des travaux de langue française sur l'évaluation formative auraient besoin d'un ancrage empirique nettement plus solide » (p. 283).

21 Un piège est de proposer des modèles théoriques séduisants mais qui prennent insuffisamment en compte les réalités contextuelles du terrain. Les propositions prescriptives ne manquent pas non plus dans la littérature en évaluation : comment « bien » évaluer, comment rendre l'évaluation plus juste et plus efficace, quels outils utiliser, quelles méthodes privilégier et appliquer ? Sans nier l'utilité de ce type de questions, notre questionnement scientifique s'intéresse, quant à lui, aux pratiques d'évaluation telles qu'elles se développent concrètement dans des situations et contextes institutionnels et culturels donnés. Ces derniers sont vus comme co-constitutifs des pratiques elles-mêmes (Lave, 1988). Notre position est que ce travail doit alors être fait *avec les acteurs de terrain*, par exemple dans le cadre de démarches de recherche dites *participatives*. Cette partie de l'article réfléchit à cette idée, en présentant les choix qui ont présidé à trois dispositifs de recherche associés à des enjeux de développement professionnel. Pour rappel, notre but ici n'est pas d'exposer les résultats de ces recherches, mais de donner à voir la nature de l'engagement des enseignants-participants par rapport à la problématique de l'évaluation des apprentissages en classe.

3.1 Penser le changement avec les acteurs de terrain dans le cadre de recherches collaboratives

22 Selon Bergold et Thomas (2012), les recherches participatives ont pour but de contribuer à changer la réalité sociale tout en tentant de comprendre les pratiques quotidiennes par la collaboration entre les différentes parties concernées. Ce courant de recherche invite à penser, avec les acteurs concernés, les potentialités de transformation des pratiques au regard notamment de leur caractère situé, de leurs enjeux, logiques, significations, actions qu'elles entraînent. Il part de :

> la conviction que le savoir n'est jamais un objet séparé de la personne, qu'il s'inscrit au contraire à même sa relation avec la personne et son expérience immédiate des circonstances dans lesquelles cette relation se produit […] la recherche participative constitue une action de la pensée sur soi, sur le groupe et sur l'environnement puisque c'est dans et par l'échange (souligné par les auteurs) que chacun apprend à accueillir les limites inhérentes à ses propres structures de signification (Bateson, 1973). (Bourassa, Bélair & Chevalier, 2007, pp. 1-2)

23 Les démarches participatives sont plurielles. Certaines allient explicitement des visées académiques et de développement professionnel pour les participants. Desgagné et ses collègues (Desgagné, 1997 ; Desgagné, Bednarz, Lebuis, Poirier & Couture, 2001) parlent, dans ce cas, de *recherche collaborative*. Les trois recherches que nous présentons ci-après ont choisi de s'inscrire dans cette conception pour les raisons principales suivantes : associer recherche et développement professionnel implique la création d'espaces de rencontres et de médiation entre chercheurs et acteurs de terrain, sous forme, le plus souvent, de réunions collectives articulées à des recueils de données en situation. Les traces (orales, écrites, non verbales) produites au cours de ces rencontres représentent des données de recherche à part entière. Les objectifs de ces espaces de rencontre sont notamment de co-définir avec les acteurs de terrain « ce qui fait problème » à propos de leurs pratiques pédagogiques, ici évaluatives, et de traiter la situation problème par des pratiques de recherche dans une visée transformative (Mottier Lopez, 2015a). Un enjeu est de construire des conditions propices au pouvoir d'agir (*empowerment*) des enseignants sur leurs propres pratiques et contextes, au regard d'un projet commun, négocié collectivement et objectivé par des démarches rigoureuses de recherche. Dans une logique d'expertise distribuée et de relation réciproque (St-Arnaud, 2003), les connaissances des acteurs de terrain sont pleinement reconnues à propos de leurs contextes et réalités professionnelles ; le chercheur apporte des connaissances scientifiques et outils méthodologiques utiles au projet collectif. Aucun partenaire ne se substitue à l'autre, mais, en termes de *process*, tous sont co-chercheurs (dans le sens d'être « en recherche sur ») par rapport à l'objet de préoccupation, partageant des questions communes mais en gardant, chacun, ses propres visées professionnelles.

24 Au-delà de ces principes épistémiques, les recensions d'écrits disponibles sur les recherches participatives montrent des formes très variées de participation, de rôles, de relations entre acteurs de terrain et chercheurs, de cadres de références convoqués, de langages, de significations négociées (e.g, Bourassa et al., 2007 ; Vinatier, Filliettaz & Kahn, 2012). Si l'on suit Callon et Latour (1991), ce sont précisément ces éléments, en termes de processus de médiation, qui seraient au cœur de l'émergence de la transformation des pratiques, plus que les propriétés mêmes du changement visé. Nous poursuivons en interrogeant quelques-uns de ces éléments dans les trois expériences de recherche participative objectivées ci-dessous.

3.2 Trois exemples de recherche : quelle relation de collaboration autour de la pratique évaluative concernée ?

25 Les trois recherches portent sur des pratiques d'évaluation des apprentissages des élèves du cycle moyen de l'école primaire genevoise (élèves de 8 à 12 ans). Toutes trois s'inscrivent dans une approche de recherche compréhensive et interprétative. Elles assument conjointement une visée de développement professionnel pour les enseignants impliqués. L'institution scolaire a reconnu la participation de ces derniers comme une formation continue à part entière. Le tableau 1 présente l'offre de formation annoncée par les chercheurs, puis les objets de préoccupation communs tels qu'ils ont été négociés (au sens interactionniste du terme) entre

les participants. Les traits caractéristiques de chaque recherche sont également présentés dans ce tableau, nous permettant de procéder directement à une discussion comparative sur les points suivants : (1) les objets d'étude, discutés au regard du contexte socio-historique présenté en début d'article ; (2) la forme d'engagement des enseignants dans les processus de recherche, en termes de rôles et responsabilités, et les principaux plans sur lesquels il paraissait possible d'agir à des fins transformatives. Notons que ce sont des recherches que nous avons menées ou dirigées. L'analyse comparative s'appuie sur notre connaissance de ces recherches, mais également sur un ensemble d'écrits qui les ont objectivées (soumission des projets et descriptifs auprès des partenaires) et de publications scientifiques que nous citerons au fil de l'article.

3.2.1 Objets de préoccupation étudiés : mise en perspective avec le contexte socio-historique de l'école primaire genevoise

26 À propos de la première recherche, le thème qui a été proposé par le chercheur à l'institution scolaire et aux enseignants au printemps 2007 portait sur « les pratiques de régulation et de différenciation en classe ». Le chercheur a choisi de ne pas directement mentionner l'évaluation formative auprès des partenaires sociaux, le contexte s'y prêtant peu de l'avis des personnes consultées. À peine une année après la réintroduction des notes à l'école primaire genevoise, toute évaluation qualitative apparaissait, d'une part, suspecte en raison notamment des confusions décrites dans la première partie de l'article. D'autre part, un ensemble d'enseignants se disait démotivé à poursuivre un projet qui, selon eux, avait avorté. Proposer le thème de la recherche sous le vocable de « régulation » permettait de redonner une place à l'évaluation formative mais sans avoir à la nommer trop explicitement.

27 L'intention était de mener une recherche longitudinale sur les processus de (auto)régulation et pratiques de différenciation pédagogique qui demandent à recueillir de l'information par des démarches formelles et/ou informelles d'évaluation formative. À partir du thème général annoncé par le chercheur, le collectif de recherche a décidé de s'interroger plus spécialement sur les interventions de l'enseignant à visée de régulation auprès des élèves dans des tâches complexes en mathématiques. L'analyse de contenu des échanges pendant les réunions collectives montre que le terme « évaluation formative » est tout de suite apparu dans les discours, associé aux termes de « régulation », « autoévaluation », « métacognition », « différenciation », notamment. Les enseignants – dont il faut dire que la plupart étaient défavorables au retour des notes, certains ayant été très investis dans la réforme scolaire – n'ont pas cherché à étudier explicitement les tensions entre leurs interventions à visée de régulation formative de l'apprentissage des élèves et la contrainte de devoir attribuer des notes certificatives. C'était comme s'il n'y avait pas/plus d'espace de problématisation pour cette question dans le contexte social de l'époque. La recherche est restée centrée sur la nature des interventions formatives des enseignants, particulièrement dans des situations dans lesquelles ils estimaient avoir des difficultés à soutenir l'autorégulation de l'élève.

28 Tableau 1 : Traits caractéristiques des trois recherches sur les pratiques d'évaluation au cycle moyen de l'école primaire genevoise (élèves 8-12 ans)

	Recherche 1 (R1)	Recherche 2 (R2)	Recherche 3 (R3)
Objet	*Offre* : Régulation et différenciation en classe *Objet négocié* : Régulation interactive entre l'enseignant et ses élèves dans des tâches complexes à des fins d'évaluation formative	*Offre* : Évaluation formative et travail de groupes *Objet négocié* : Principes organisateurs (heuristiques) susceptibles de relever d'une « évaluation collaborative » entre élèves et avec l'enseignant	*Offre* : Évaluation certificative à l'école primaire genevoise dans des tâches complexes *Objet négocié* : Repères communs pour élaborer les épreuves et évaluer les productions des élèves
Partenaires	1 chercheur 9 enseignants	2 chercheurs 6 enseignants	5 chercheurs 29 enseignants, en 2 cohortes
Durée	3 années scolaires (2007-2010) 6 rencontres / année (total, 126 heures de réunions collectives)	1 année scolaire (2013-2014) 4 rencontres (total, 20 heures)	5 semestres, 2012-2014 3 rencontres pour chaque cohorte (total, 17 heures pour chaque cohorte)
En plus des données « réunions collectives »	Données recueillies par chaque enseignant dans sa classe + écrits réflexifs individuel	Données recueillies par un des chercheurs dans chaque classe + entretiens semi-dirigés avec enseignant(s)	Données recueillies par les chercheurs auprès de chaque enseignant + entretiens semi-dirigés
Négociation de significations collectives	À partir des expériences en classe + données d'observation : - Co-analyses et co-interprétations des données recueillies sur les pratiques o à partir d'outils de la recherche o croisés aux savoirs professionnels des enseignants	À partir des expériences en classe + données d'observation : - Échanges sur les conditions, outils, situations favorables à une évaluation collaborative entre élèves et avec l'enseignant - Dans des activités scolaires réalisées en petits groupes d'élèves	À partir des expériences en classe + données d'observation : - Confrontation des pratiques évaluatives des enseignants (construction des épreuves + appréciation des réponses des élèves) - Construction de consensus et reconnaissance de divergences
Approches de recherche	Recherche collaborative	Recherche action collaborative	Recherche collaborative à composante délibérative
Références	*Mottier Lopez (2012)*	*Morales Villabona (2013)*	*Mottier Lopez, Tessaro, Dechamboux & Morales Villabona (2012)*

Le deuxième cas est une recherche doctorale, en cours, de Morales Villabona (2013). Six ans ont passé depuis la réintroduction des notes à Genève. Une analyse systématique des offres proposées par les « catalogues[8] » de formation continue des enseignants de l'école primaire genevoise a montré que, depuis, plus une seule de ces offres n'a porté directement sur l'évaluation formative, contrairement aux années qui ont précédé la réforme scolaire, ainsi que pendant sa phase de généralisation (Mottier Lopez, 2014). Comme dit plus haut, si l'évaluation formative reste inscrite dans le règlement d'application de la loi scolaire, elle n'apparaît plus explicitement comme un objet qui justifie des interventions ciblées dans les programmes de formation continue. Morales Villabona a néanmoins tenté de proposer une offre sur le thème de « l'évaluation et le travail collaboratif en groupes d'élève ». Le descriptif annonçait explicitement l'évaluation formative comme objet d'étude, problématisé par rapport au travail en petits groupes d'élèves, une modalité très présente dans les moyens d'enseignement utilisés par les enseignants genevois. Le descriptif proposait également de réfléchir aux liens entre évaluations formative et certificative à partir des pratiques des participants et d'une démarche de recherche menée collectivement.

29 Ce thème n'a pas remporté suffisamment de succès pour qu'une session de formation s'ouvre (exigence d'au moins 10 enseignants intéressés sur base de volontariat). Cela peut s'interpréter comme un désintérêt persistant vis-à-vis de l'évaluation formative. De façon moins radicale, on peut penser que ce thème ne semble pas suffisamment attrayant par rapport aux autres offres, ou encore que la dimension « recherche » peut avoir représenté un frein.

30 Les enseignants ont donc été contactés personnellement. Deux enseignants de la première recherche ont répondu positivement et ils ont « recruté » trois de leurs collègues faisant partie du même établissement scolaire, avec le soutien de leur directeur d'école. Une sixième personne, d'un autre établissement scolaire, a demandé à participer en réponse à l'offre de formation continue. Le dispositif de recherche impliquait un accompagnement des enseignants afin d'explorer, avec eux, l'idée d'une « évaluation collaborative » entre élèves et avec l'enseignant, dans le prolongement de propositions de la littérature anglophone (e.g., Hargreaves, 2007) mais peu documentées en contexte réel de classe. En alternance avec des expériences en classe, les temps collectifs[9] de cette recherche visaient à construire avec les enseignants un *cadre heuristique pour agir* (non prescriptif) pour penser les pratiques évaluatives et des possibles transformations. L'analyse de contenu des réunions collectives dans Mottier Lopez et Morales Villabona (soumis) montre que le collectif de recherche a élaboré des principes heuristiques essentiellement en termes de rôles partagés dans la relation évaluative entre élèves et avec l'enseignant, ainsi que sur un plan didactique dans la structuration de l'activité (productions écrites en petits groupes) et des critères d'évaluation. Par rapport à la première recherche, on retiendra que, dans ce cas-ci, l'évaluation formative est formelle, instrumentée par différents outils conçus par les enseignants. Dans la première

recherche, l'évaluation formative était avant tout processuelle. Ici, l'évaluation se nomme, s'explicite, s'observe dans des traces d'appréciation critériée produites par les élèves, qui se sont auto-évalués dans des travaux de groupe, puis réexploitées par l'enseignant avec toute la classe. Les enjeux de l'évaluation certificative, sans être exclus des préoccupations du collectif de recherche, n'ont pas représenté un objet explicite de problématisation pour les enseignants. Dans leur représentation, l'évaluation formative, encore plus si elle implique des interactions entre élèves, reste principalement attachée à des visées d'enseignement et d'apprentissage. Elle apparaît décrochée, en partie, des démarches de certification qui convoquent principalement des modalités de travail individuel de la part des élèves.

31 La troisième recherche[10], dont le dispositif est décrit dans Mottier Lopez, Tessaro, Dechamboux et Morales Villabona (2012), change volontairement de perspective. Puisque l'évaluation certificative est clairement sur le devant de la scène politico-pédagogique, bien que peu pensée par rapport aux pratiques d'évaluation formative, l'option est qu'elle devienne l'objet premier d'étude. Cette recherche a nécessité des discussions préalables avec plusieurs cadres de l'institution scolaire et les directeurs d'école pour obtenir les autorisations d'usage et parvenir à réunir une trentaine d'enseignants. L'institution a souhaité que l'étude porte sur les disciplines scolaires du français et des mathématiques (initialement, le projet était de contraster les observations entre une discipline qui a du poids dans les décisions pronostiques de fin d'école primaire et une discipline « mineure »). Il nous a également été demandé de prendre en compte, dans nos échanges avec les enseignants, un nouveau plan d'études en cours d'implantation (le plan d'études romand – PER), objet sensible d'une politique inter-cantonale d'harmonisation de la scolarité obligatoire en Suisse. Au plan symbolique, ces quelques négociations montrent combien l'évaluation certificative est sous contrôle de l'institution scolaire qui définit les règles du jeu. Aucune transaction de ce type n'a été nécessaire pour les deux autres recherches sur l'évaluation formative, dont le cadre de prescription institutionnelle est plus souple.

32 Une fois un accord trouvé, les chercheurs ont annoncé aux enseignants que la recherche visait à mieux comprendre leurs pratiques d'évaluation certificative dans leur contexte d'exercice quotidien. Le dispositif proposé alternait des phases d'observation de pratiques dans le terrain et des réunions collectives pour expérimenter des pratiques dites de *modération sociale*. Considérée comme une pratique à part entière de développement professionnel dans certains contextes anglophones, la modération sociale a pour but de co-construire des repères partagés et des consensus pour fonder la validité des jugements évaluatifs situés des enseignants (Adie, Klenowski & Wyatt-Smith, 2012). Dans la recherche de Mottier Lopez et collaborateurs, les échanges de modération sociale ont porté sur deux aspects en particulier : (1) la construction de l'outil d'évaluation, (2) l'appréciation du travail de l'élève sur des tâches données. Au plan scientifique, un des objectifs était de mieux comprendre les contraintes qui pèsent sur les pratiques situées et d'observer la façon dont les enseignants « jouent » avec des processus formatifs au sein même de l'évaluation certificative, notamment en termes d'ajustements comme observé dans Mottier Lopez et al. (2012).

3.2.2 Engagement des enseignants dans les processus de recherche et visées de changement poursuivies

33 Dans une approche de recherche collaborative, les enseignants sont accompagnés à développer une posture d'être en recherche sur leurs pratiques et, plus largement, d'une communauté de pratique (au sens de Wenger, 1998) à laquelle ils appartiennent. Un point déterminant, selon Bergold et Thomas (2012), serait leur implication dans les prises de décision et responsabilité partagée vis-à-vis du projet commun. C'est pourquoi, pour chacune des trois recherches qui nous intéressent, nous interrogeons ci-après l'engagement des enseignants sur deux plans : (1) à propos des pratiques en classe étudiées, (2) et de la conception du dispositif de recherche et de sa mise en œuvre. En termes de développement professionnel, nous discutons ensuite les dimensions sur lesquelles il semblait possible d'agir par le moyen de l'engagement sollicité.

34 Le tableau 2 résume la tendance des trois recherches du point de vue de la *responsabilité* des enseignants dans les *prises de décision*. Au niveau des pratiques étudiées dans leurs classes, leur responsabilité était maximale dans les trois recherches. A aucun moment, le chercheur ne

s'est substitué à leur expertise. Ce sont toujours les enseignants qui ont décidé les dispositifs à expérimenter à partir de leurs pratiques existantes (R1 et R3) ou en lien avec des démarches innovantes (R2). Dans les deux premières recherches, des concertations ont eu lieu dans le collectif de recherche, puis chaque enseignant conservait la liberté de ses choix effectifs à titre individuel. Dans la deuxième recherche, les enseignants ont choisi, en plus, de se concerter entre eux hors des réunions collectives. Quant à la troisième recherche, les pratiques étudiées ont été observées deux fois en amont des modérations sociales, puis encore deux fois après. À chaque fois, ce sont les enseignants (ou parfois duos d'enseignants appartenant à un même établissement scolaire) qui ont décidé ce qu'ils donnaient à voir aux chercheurs. Ci-après, nous commentons ces différents cas de figure au regard de l'engagement des enseignants dans les processus de recherche.

35 Tableau 2 : Responsabilité des enseignants dans les prises de décision

	Au niveau des pratiques étudiées en classe	**Au niveau du dispositif de recherche**
R1	*Maximale* Collectif de recherche - individuellement	*Maximale* Dans toutes les étapes de la recherche
R2	*Maximale* Collectif de recherche – équipe autonome d'enseignants impliqués dans la recherche – individuellement	*Partielle* Orientation de la problématique, choix du dispositif de recueil de données, en partie co-analyse et co-interprétation
R3	*Maximale* Individuellement (rares duos d'enseignants) – Collectif de recherche – Individuellement (duos)	*Minimale* Choix concertés pour certains éléments du dispositif de recueil de données, validité de signifiance

36 Dans la première recherche, les enseignants étaient co-responsables de l'ensemble des choix méthodologiques. Ils ont contribué à toutes les phrases de la recherche : définition de la problématique et questions de recherche, conception du dispositif, recueil des données, analyses et communications dans des colloques et dans leur milieu professionnel, co-écriture d'un article (Mottier Lopez et al., 2010). Quant aux activités didactiques étudiées en classe, elles ont été choisies par le collectif de recherche à partir des expériences et connaissances des enseignants, ainsi que sur la base d'analyses a priori. Le groupe a décidé que chaque enseignant était responsable de recueillir dans sa propre classe des extraits d'interactions qu'il jugeait représentatifs d'une « régulation interactive » entre lui et un / plusieurs de ses élèves. Chaque enseignant devait ensuite transcrire les extraits qu'il avait retenus et rédiger une réflexion critique avant d'apporter ce matériel en réunion, notamment pour expliquer ses choix et contextualiser les données apportées en réunion. Le postulat était que les choix effectués représentaient une source d'information significative sur ce que représente, pour la profession, une évaluation formative interactive, visant à soutenir l'autorégulation d'un élève dans une tâche complexe. Le chercheur ne s'est donc jamais déplacé dans les classes.

37 En termes de développement professionnel, les dimensions sur lesquelles il semblait pouvoir possible agir par le moyen de cette recherche, portaient essentiellement sur la nature des guidages interactifs à visée de régulation de l'enseignant. Autrement dit, l'objet concernait directement les enseignants par rapport à leurs possibilités et compétences d'intervention auprès de leurs élèves en situation de classe. Les interprétations collectives, fondées sur des unités d'analyse définies par la littérature scientifique[11], croisées aux connaissances professionnelles des enseignants, ont pris en compte un ensemble de contraintes telles les enjeux didactiques et d'apprentissage des élèves dans des temporalités courtes et longues, l'idéologie des moyens d'enseignement utilisés et, plus largement, des courants actuels qui tendent à encourager le développement de compétences chez les élèves dans des situations complexes. Cette recherche a duré trois ans, articulée autour de deux cycles de recueil de données, d'analyses et interprétations collectives, visant une possible transformation d'un cycle à l'autre.

38 Passons à la recherche doctorale de Morales Villabona qui a impliqué les enseignants dans les choix relevant du processus de recherche mais de façon moins soutenue que dans le cas précédent. Les enseignants ont été sollicités dans la définition de ce que pourrait être une « évaluation collaborative » dans leur classe. Ils ont largement contribué aux choix des dimensions à problématiser qui faisaient sens pour eux. Ils ont essentiellement questionné les apprentissages susceptibles d'être évalués par cette forme d'évaluation (compétences sociales de collaboration, compétences disciplinaires), la co-construction avec les élèves des critères d'évaluation, l'exploitation des évaluations de groupe à des fins de régulation et co-régulation (Mottier Lopez & Morales Villabona, soumis).

39 Les enseignants étaient également responsables d'envisager les grandes lignes de ce qu'ils pourraient expérimenter dans leur classe à partir du cadre commun co-élaboré. Mais contrairement à la première recherche, les enseignants n'avaient pas à recueillir des données empiriques, ni à les mettre en forme. C'est le chercheur qui s'est rendu dans chaque classe, a géré les observations et mené des entretiens post-observatoires. Son rôle était de formaliser les données d'observation pour proposer un matériau à interpréter collectivement en réunion. Pour ce faire, il a préparé un dossier composé de plusieurs pièces : description de chaque activité en classe (notamment, discipline scolaire, objectifs disciplinaires et/ou transversaux visés, organisation sociale de la tâche, de l'évaluation en petits groupes d'élèves, temporalités, questions en suspens), extraits transcrits d'interactions entre élèves, outils d'évaluation élaborés. À partir de ce dossier, les enseignants étaient engagés à mener des analyses réflexives et critiques avec les chercheurs. Comparativement à la première recherche qui s'était fondée, entre autres, sur des outils d'analyse et de codage issus de la littérature scientifique, dans ce cas-ci, les analyses essentiellement dialogiques se sont appuyées uniquement sur des concepts propres au champ de l'évaluation et les connaissances professionnelles des enseignants. Elles n'ont pas débouché sur un codage du matériau par exemple.

40 Les dimensions sur lesquelles une transformation était visée portaient non pas sur les interventions directes des enseignants auprès de leurs élèves (R1), mais sur l'idée de pouvoir impliquer différemment les élèves dans une « approche sociale » de l'autoévaluation (Van Gennip, Segers & Tillema, 2009). Évidemment, l'enseignant est actif dans cette intention, entre autres par la conception des situations didactiques et des outils d'autoévaluation, par la négociation avec les élèves des significations associées à la démarche évaluative et de son usage. Le défi pour le collectif de recherche était alors de penser le développement professionnel des enseignants à la lumière essentiellement des observations et informations *produites par les élèves* en situation de classe. Sur le temps d'un semestre, des va-et-vient entre réunions collectives et expériences en classe ont été réalisés pour affiner les questionnements et pour envisager des ajustements successifs des pratiques en classe. En ce sens, ce dispositif se rapproche d'une recherche-action ou recherche-développement, tout en assumant une dimension collaborative comportant une visée de développement professionnel (Bourassa et al., 2007).

41 Enfin, dans la recherche 3, l'ensemble du dispositif méthodologique était géré par les chercheurs[12]. Quelques éléments seulement étaient sous la responsabilité des enseignants, dont le choix des contenus disciplinaires évalués en français et en mathématiques, qui devaient être communs entre les classes à des fins de comparabilité, et la temporalité des observations dans le terrain. Les enseignants n'ont pas eu à recueillir des données empiriques, ni à concevoir des dispositifs en classe au regard d'un questionnement collectivement élaboré. Dans ce cas-ci, l'engagement des enseignants était d'un autre ordre, associé aux pratiques de modération sociale expérimentées au cours de trois rencontres (voir tableau 1). À partir de données d'observation antérieures à ces modérations sociales, mises en forme par les chercheurs, les enseignants devaient confronter leurs pratiques évaluatives au regard d'épreuves et productions concrètes d'élèves. Ils avaient alors la responsabilité de construire des *consensus* sur les dimensions de l'évaluation certificative (pour rappel, autour de la construction de l'épreuve et de l'appréciation de la production de l'élève) en distinguant trois niveaux, adaptés de Morrissette (2011) :

1. Les dimensions vues comme partagées et devant être communes entre eux ;
2. Les dimensions vues comme pouvant être légitimement différentes. Ces différences sont admises et acceptables selon eux ;
3. Les dimensions pour lesquelles il y a un désaccord social. Les différences de pratiques sont contestées par le collectif.

42 Tout en valorisant des processus collectifs de construction de significations, cette troisième recherche a pour particularité de valoriser la confrontation de points de vue et la construction de compréhensions vue comme reconnues et partagées à propos d'accords et de désaccords sociaux. Les points de vue demandent alors à être argumentés et à être compris par le collectif de recherche. Burchardt (2014) parle, dans ce cas, de recherche délibérative (*deliberate research*). Cet auteur insiste également sur les apports d'experts externes afin de décentrer les débats et offrir d'autres cadres aux processus de délibération. Le dispositif de la troisième recherche a joué sur cette décentration, en organisant une alternance entre des moments de confrontation en petits groupes d'enseignants hors de la présence des chercheurs, puis collectivement avec eux (pour plus de détail, voir la description dans Mottier Lopez et al., 2012). Entre chaque jour de modération sociale (espacé de quinze jours), les chercheurs ont analysé les consensus émergeant collectivement à partir des traces écrites produites pendant les séances. Puis, au début de chaque nouvelle réunion, ils les ont soumis au groupe. La responsabilité des enseignants était alors de confirmer, ajuster, réfuter si besoin les propositions des chercheurs dans une logique de *validité de signifiance*[13] (Pourtois & Desmet, 1997). Autrement dit, l'engagement des enseignants était de contribuer activement à faire émerger, toujours dans une approche délibérative, les éléments d'une culture partagée de l'évaluation, ainsi que les singularités, plus ou moins légitimes, susceptibles de déboucher sur des variations de jugements évaluatifs entre enseignants. En termes de recherche, cet engagement consistait aussi à être co-responsables de la validation des éléments dégagés.

43 Les dimensions sur lesquelles cette recherche visait à agir portaient essentiellement sur les valeurs, les connaissances professionnelles, les cadres de référence convoqués plus ou moins explicitement par les enseignants dans leurs pratiques évaluatives. Elle visait plus une réflexion collective et critique sur les raisonnements évaluatifs et leurs conséquences qu'une action directe sur les pratiques d'évaluation elles-mêmes. C'est ainsi que ces espaces ont fait émerger plusieurs tensions vécues par les enseignants entre leurs interventions formatives à des fins de soutien d'apprentissage et les exigences de l'évaluation certificative, donnant à voir des dilemmes ou ruptures situationnelles entre les différents fonctions de l'évaluation. Les observations post-modérations sociales, en cours d'analyse, montreront dans quelle mesure des transformations se sont opérées.

4. Éléments de conclusion

44 La mise en perspective des trois recherches que nous avons exposées illustre, en partie, la pluralité possible des modalités d'engagement des acteurs de terrain dans des recherches participatives. Pour rappel, celles-ci assument pleinement la finalité de contribuer à un changement social par la collaboration avec les acteurs concernés, en tentant de comprendre avec eux leurs pratiques situées (Bergold & Thomas, 2012). Parmi les différentes formes de recherche participative, nous avons privilégié des dispositifs qui revendiquent le double but de soutenir des processus de développement professionnel tout en contribuant à produire des savoirs scientifiques sur les pratiques étudiées (Desgagné et al., 1997, 2001). Ces recherche dite *collaboratives* peuvent être de différents ordres, certaines assumant également une visée de recherche-action comme c'est le cas de la recherche de Morales Villabona, d'autres s'appuyant sur une composante délibérative au sens de Burchardt (2014) comme dans la recherche de Mottier Lopez et al. (2012).

45 Pour les chercheurs et enseignants, les trois recherches avaient pour but commun de tenter de mieux comprendre les pratiques d'évaluation des apprentissages en contexte de classe, par rapport à des enjeux pédagogico-didactiques et sociaux plus larges, puis pour les enseignants de pouvoir agir sur leurs pratiques en toute responsabilité. L'hypothèse est que les croyances et les valeurs des participants sont susceptibles de se transformer par leur engagement dans

des processus de recherche. Mais il s'agit alors d'une participation qui se pense et s'organise, notamment en termes de responsabilités et prises de décision vis-à-vis des pratiques en classe et du dispositif de la recherche et de sa mise en œuvre, comme nous avons tenté de le montrer ?

46 Les fonctions formative et certificative de l'évaluation sont souvent considérées comme irréductiblement opposées, notamment quand des enjeux de sélection s'en mêlent (e.g., Perrenoud, 1998). Notre proposition, développée dans Mottier Lopez (2015c), est plutôt de les conceptualiser, malgré leurs oppositions, comme étant constitutives d'un même système. Nous suivons Morin (1977) quand il dit qu'il ne s'agit pas d'opposer mais de distinguer et de lier les tendances duales. Cela demande d'une part d'intégrer les caractéristiques associées à chaque élément et, d'autre part, de faire émerger des caractéristiques propres à l'articulation et à l'interaction entre les éléments. Évidemment, les recherches que nous avons présentées sont modestes par rapport à cet enjeu. Plus généralement, peu d'études francophones ont été entreprises sur les relations, tensions, ruptures, recherche d'alignement entre les différentes fonctions de l'évaluation qui, pourtant, doivent coexister au sein d'une même classe et d'un même système scolaire. Les recherches que nous avons présentées s'inscrivent dans cette visée, tentant de dépasser une approche strictement instrumentale souvent associée à l'évaluation scolaire, afin d'appréhender les processus, les tensions, les significations, les enjeux multiples tels qu'ils sont perçus par les enseignants, en relation avec le contexte socio-historique des pratiques et à la fois les contextes singuliers qui se négocient au quotidien dans les classes. Les questions que cette visée soulève sont vives et méritent de poursuivre les recherches.

Bibliographie

Abernot, Y. (2013). Evolution d'homo evaluatus. In J.-F. Marcel & H. Savy, (Ed.), *Évaluons, évoluons. L'enseignement agricole en action* (pp. 185-196). Dijon, France : Educagri éditions.

Adie, L., Klenowski, V. & Wyatt-Smith, C. (2012). Towards an understanding of teacher judgement in the context of social moderation. *Educational Review, 64*(2), 223–240.

Allal, L. (2006). La fonction de régulation de l'évaluation : constructions théoriques et limites empiriques. In G. Figari & L. Mottier Lopez (Ed*.), Recherche sur l'évaluation en éducation : problématiques, méthodologies et épistémologie* (pp. 223-230). Paris: L'Harmattan.

Allal, L. & Mottier Lopez, L. (2005). Formative Assessment of Learning: A Review of Publications in French (pp. 241-264). In *Formative Assessment - Improving Learning in Secondary Classrooms.* Paris: OECD Publication.

Allal, L., Bain, D. & Perrenoud, P. (Ed.) (1993). *Évaluation formative et didactique du français.* Lausanne : Delachaux et Niestlé.

Allal, L., Cardinet, J. & Perrenoud, P. (Ed.). (1979). *L'évaluation formative dans un enseignement différencié.* Berne : Peter Lang.

Allal, L., Bétrix Köhler, D., Rieben, L., Rouiller, Y., Saada-Robert, M. & Wegmuller, E. (2001). *Apprendre l'orthographe en produisant des textes.* Fribourg: Éditions Universitaires.

Bergold, J. & Thomas, S. (2012). Participatory Research Methods: A Methodological Approach in Motion [110 paragraphs]. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 13 (1). http://nbn-resolving.de/urn:nbn:de:0114-fqs1201304

Blanchet, A. (2010). La réforme à l'épreuve de l'évaluation dans le canton de Vaud. In P. Gilliéron Giroud & L. Ntamakiliro (Ed.), *Réformer l'évaluation scolaire : mission impossible ?* (pp. 79-112). Berne : Peter Lang.

Bloom, B. S., Hasting, J. T. & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning.* New York: McGraw-Hill.

Bourassa, M., Bélair, L.M. & Chevalier, J. (2007). Les outils de la recherche participative. *Éducation et francophonie, 23* (2), 1-11.

Burchardt, T. (2014). Deliberative research as a tool to make value judgements. *Qualitative Research, 14* (3), 353-370.

Callon, M. & Latour, B. (Ed.) (1991). *La science telle qu'elle se fait.* Paris : La Découverte.

Cardinet, J. (1983) *Des instruments d'évaluation pour chaque fonction.* Neuchâtel : Institut romand de recherches et de documentation pédagogiques.

Desgagné, S. (1997). Le concept de recherche collaborative : l'idée d'un rapprochement entre chercheurs universitaires et praticiens enseignants. *Revue des sciences de l'éducation, 23* (2), 371-393.

Desgagné, S., Bednarz, N., Lebuis, P., Poirier, L. & Couture, C. (2001). L'approche collaborative de recherche en éducation : un rapport nouveau à établir entre recherche et formation. *Revue des sciences de l'éducation, 27* (1), 33-63.

Felouzis, G. & Hanhart, S. (2011). *Gouverner l'éducation par les nombres ? Usages, débats et controverses.* Bruxelles : De Boeck (collection Raisons éducatives).

Hargreaves, E. (2007). The validity of collaborative assessment for learning. *Assessment in Education: Principles, Policy and Practice, 14* (2), 185-199.

Gilliéron Giroud, P. & Ntamakiliro, L. (Ed.) (2010*). Réformer l'évaluation scolaire : mission impossible ?* Berne: Peter Lang.

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life.* Cambridge, MA : Cambridge University Press.

Morales Villabona, F. (2013). *Modélisation d'une évaluation collaborative des apprentissages des élèves : alternance entre un dispositif de formation et des pratiques en classe à l'école primaire.* Canevas de thèse en Sciences de l'éducation, Université de Genève, Suisse.

Morin, E. (1977). *La méthode* (Tome I : La nature de la nature). Paris : Le seuil.

Morrissette, J. (2011). Vers un cadre d'analyse interactionniste des pratiques professionnelles. *Recherches qualitatives, 30* (1), 38-59.

Mottier Lopez, L. (2012). Interprétation d'une recherche collaborative dans les termes de la « valuation » de Dewey. *Travail et apprentissage, 9*, 183-199.

Mottier Lopez, L. (2014). L'évaluation pédagogique va-t-elle enfin marcher sur ses deux pieds ? Les enseignements de l'histoire récente de l'école primaire genevoise. In D. Laveault (Ed.), « Les politiques d'évaluation en éducation. Et après ? » *Éducation et francophonie, XLII* (3), 85-101. http://www.acelf.ca/c/revue/pdf/EF42-3-NoSpecial-085-MOTTIER_LOPEZ.pdf

Mottier Lopez, L. (2015, janvier). *Cadres psychologiques et didactiques : entre rapprochements et oppositions pour contribuer à la modélisation de l'évaluation formative des apprentissages des élèves*. Communication au 26ème colloque international de l'ADME-Europe, Université de Liège, Belgique.

Mottier Lopez, L, (2015a). Au cœur du développement professionnel des enseignants, la conscientisation critique. Exemple d'une recherche collaborative sur l'évaluation formative à l'école primaire genevoise. *Carrefour de l'éducation, 39,* 119-135.

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). *Évaluer et apprendre en classe. Enjeux des évaluations formative et certificative.* Bruxelles : De Boeck (Le Point sur… Pédagogie).

Mottier Lopez, L. & Laveault, D. (2008). L'évaluation des apprentissages en contexte scolaire : développements, enjeux et controverses. *Mesure et évaluation en éducation, 3,* 5-34.

Mottier Lopez, L., Borloz, S., Grimm, K., Gros, B., Herbert, C., Methenitis, J., Payot, C., Pot, M., Schneuwly, J.-L. & Zbinden, S. (2010). Les interactions de la régulation entre l'enseignant et ses élèves. Expérience d'une recherche collaborative. In L. Lafortune, S. Fréchette, N. Sorin, P.-A. Doudin & O. Albanese (Ed.), *Approches affectives, métacognitives et cognitives de la compréhension* (pp. 33-50). Québec : Presses de l'Université du Québec, collection éducation-intervention.

Mottier Lopez, L., Tessaro, W., Dechamboux, L. & Morales Villabona, F. (2012). La modération sociale : un dispositif soutenant l'émergence de savoirs négociés sur l'évaluation certificative des apprentissages des élèves. *Questions vives, 6* (18). http://questionsvives.revues.org/1235

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

Perrenoud, P. (2001). Évaluation formative et évaluation certificative : postures contradictoires ou complémentaires ? *Formation professionnelle suisse, 4*, 25-28.

Salamin, J.-P. (Ed.). (1986). *École primaire suisse : une école pour apprendre : 22 thèses pour le développement de l'école primaire : projet SIPRI.* Berne : CDIP.

St-Arnaud, Y. (2003). *L'interaction professionnelle : efficacité et coopération.* Montréal : Les presses de l'université de Montréal.

van Gennip, N.A., Segers, M.S. & Tillema, H.H. (2009) Peer assessment for learning from a social perspective: The influence of interpersonal variables and structural features. *Educational Research Review, 4*, 41-54.

Vinatier, I., Filliettaz, L. & Kahn, S (Ed.). (2012). Enjeux, formes et rôles des processus collaboratifs entre chercheurs et professionnels de la formation : pour quelle efficacité. *Travail et apprentissage, 9.*

Weiss, J. (1996). *Vers une conception cohérente de l'évaluation pour la scolarité obligatoire en Suisse romande et au Tessin : une évaluation pour apprendre et pour choisir : voies et moyens : rapport no 2.* Neuchâtel: IRDP.

Weiss, J., Wirthner, M., Bain, D., Greub, J. & Perrenoud, B. (1998). *Pour une évaluation plus formative : actes du Colloque « Évaluation 97 ».* Neuchâtel : Institut romand de recherches et de documentation pédagogiques.

Wenger, E. (1998). *Community of practice: Learning, meaning, and identity.* New York: Cambridge University Press.

Wiliam, D. (2011). *Embedded formative assessment.* Bloomington, IN: Solution Tree Press.

Wirthner, M. (2010). Un bilan des réformes de l'évaluation du travail des élèves en Suisse romande. In P. Gilliéron Giroud & L. Ntamakiliro, L. (Ed.), *Réformer l'évaluation scolaire : mission impossible ?* (pp. 13-34). Berne : Peter Lang.

Notes

1 À partir de la proposition de Scriven dans le cadre de l'évaluation de programmes et dispositifs de formation.

2 On notera le choix de l'auteur de parler d'évaluation certificative plutôt que d'évaluation sommative. Pour notre part, nous privilégions également le terme évaluation certificative, excepté quand nous citons un auteur qui, lui, ne l'utilise pas (par exemple Salamin, dans cet article).

3 Une rencontre informelle en 1977 à l'Université de Mons, puis trois rencontres en 1978 à l'Université de Genève.

4 L'évaluation étant alors principalement théorisée par les travaux psychométriques et docimologiques.

5 Dans Allal et Mottier Lopez (2005), nous donnons à voir les principaux points d'élargissement du modèle de l'évaluation formative dans les travaux de langue française.

6 Nous citons les degrés scolaires tels qu'ils étaient désignés à l'époque. Depuis, cette dénomination a changé en raison d'une politique d'harmonisation des structures entre les cantons. Voir : http://www.edk.ch/dyn/11737.php

7 En plus des revues et ouvrages spécialisés en évaluation, l'analyse a été élargie à des revues telles que *La lettre de l'AIRDF*, *Recherches en didactique du français*, *Revue Pratiques*, *Revue Repères*, *Recherche en didactique des mathématiques*. Nous remercions Sophie Serry pour l'aide apportée à cette recherche bibliographique.

8 Terme utilisé dans le terrain genevois pour désigner le document officiel qui regroupe, chaque année, les offres de formation continue.

9 L'animation des séances collectives étaient assurées par Morales Villabona et nous-même en tant que directrice de la thèse.

10 Cette recherche est soutenue par le Fonds national Suisse de la recherche scientifique (requête n °100013_143453/1) que nous remercions.

11 Nous renvoyons le lecteur à Mottier Lopez (2015b) qui présente en détail les outils d'analyse et leur usage.

12 Ces chercheurs sont les membres du groupe de recherche EReD (évaluation, régulation et différenciation des apprentissages dans les systèmes d'enseignement) que nous dirigeons, à savoir Lionel Dechamboux, Fernando Morales Villabona, Sophie Serry et Walther Tessaro.

13 Cette validation consiste à soumettre les interprétations des chercheurs aux acteurs concernés en vue d'une mise en discussion, considérée comme un excellent moyen de vérifier leur plausibilité et d'augmenter leur crédibilité.

Pour citer cet article

Référence électronique

Lucie Mottier Lopez, « L'évaluation formative des apprentissages des élèves : entre innovations, échecs et possibles renouveaux par des recherches participatives », *Questions Vives* [En ligne], N° 23 | 2015, mis en ligne le , consulté le 14 février 2016. URL : http://questionsvives.revues.org/1692 ; DOI : 10.4000/questionsvives.1692

À propos de l'auteur

Lucie Mottier Lopez
Professeure associée, dirige le groupe de recherche EReD (évaluation, régulation et différenciation des apprentissages dans les systèmes d'enseignement). Faculté de psychologie et des sciences de l'éducation, Université de Genève

---

Droits d'auteur

Tous droits réservés

Résumé

Entrées d'index

Mots-clés : évaluation formative, réforme scolaire, recherche collaborative, engagement des enseignants dans la recherche
Keywords : formative assessment, school reform, collaborative research, teachers' commitment in theresearch process