

D'un référentiel d'évaluation fixe à une co-constitution référentielle dynamique, ce que nous apprend le jugement situé de l'enseignant

Lucie MOTTIER LOPEZ¹ et Lionel DECHAMBOUX²

¹ Université de Genève, Faculté de psychologie et des sciences de l'éducation

² Haute École Pédagogique du canton de Vaud et Université de Genève, Suisse

Résumé

Il est aujourd'hui largement admis que l'évaluation des apprentissages des élèves par les enseignants doit s'appuyer sur un référentiel d'évaluation exprimé sous forme d'objectifs et/ou de critères d'évaluation à partir de plans d'études institués. Peu de travaux ont examiné la façon dont ce référentiel est susceptible d'être transformé au fil de l'activité évaluative de l'enseignant. Quelle relation observe-t-on entre les éléments préexistants formalisés dans le référentiel d'évaluation et d'autres éléments informels qui interviennent également dans le jugement de l'enseignant ? Quel rôle les contextes jouent-ils dans la définition de ces éléments et leur convocation parfois émergente dans les processus évaluatifs ? L'objectif de cet article est de présenter quelques éléments de modélisation autour de ces questions à partir d'études empiriques préalables et de propositions de la littérature. Les études concernent des pratiques d'évaluation certificative d'enseignants du deuxième cycle (élèves de 8 à 12 ans) de l'école primaire genevoise, en français (production écrite) et en mathématiques (résolution de problèmes). À partir de l'exposition de résultats, l'article interroge le modèle théorique de l'évaluation autour de deux hypothèses de travail présentées successivement, celle d'une relation dialectique entre référent et référé dans le jugement évaluatif en acte de l'enseignant, et celle d'une part prise par des contextes pluriels ayant des fonctions particulières dans les processus d'interprétation et de prise de décision évaluatives. L'enjeu scientifique, en proposant des éléments de modélisation, est de mieux comprendre comment référents et référés agissent dans les pratiques d'évaluation situées des enseignants.

Mots-clés

Évaluation certificative, modélisation théorique, référentialisation, référentiel d'évaluation, référé, référent.

Abstract

It is now widely accepted that the evaluation of student learning by teachers should be based on an evaluation framework expressed in terms of objectives and/or evaluation criteria based on the established curricula. Little research has examined how this framework is likely to be affected in the course of teachers' evaluative activity, in particular during the grading process. What is the relation between preexisting elements, as formalized in the evaluation framework, and other informal elements that also operate during the grading process? What role do contexts play in defining these elements and in their convocation by teachers during the grading process? The objective of this article is to present some elements of modeling about these issues, based on empirical studies and proposals from the literature. The field studies concerned classroom assessment practices of teachers of primary school in the canton of Geneva (pupils aged 8 to 12) in French (writing) and in mathematics (problem solving). From the empirical results, the article questions the theoretical model of evaluation around two working hypotheses: first, a dialectical relationship between referents and referees in teachers' evaluative judgement in action; secondly, the role played by multiple contexts in the interpretation and decision-making processes underlying teachers' evaluations. The scientific challenge is to elaborate elements of a theoretical modeling in order to better understand how referents and referees act in situated evaluation practices of teachers.

Keywords

Evaluation framework, referee, referent, referentialization, summative assessment, theoretical modeling.

1. Introduction

Il semble aujourd'hui largement admis que l'évaluation des apprentissages des élèves par les enseignants doit s'appuyer sur un référentiel d'évaluation prescrit, exprimé sous forme d'objectifs et/ou de critères d'évaluation à partir de plans d'études institués. Le référentiel désigne « un construit social qui clarifie les normes d'une activité ou un sens donné à des systèmes sociaux. Il est ce par rapport à quoi un jugement ou un sens est donné à un objet ou une action » (Cros et Raisky, 2010 : 107). Toutefois, même si de nombreux travaux ont porté sur la question du référentiel de l'évaluation des apprentissages des élèves, nos analyses systématiques de la littérature montrent que peu examinent la façon dont ce référentiel est susceptible d'être spécifié et transformé au fil de l'activité évaluative tout en infléchissant en retour le jugement produit au cours de cette activité (Allal et Mottier Lopez, 2005 ; Mottier Lopez et Laveault, 2008 ; Mottier Lopez, 2015). Peu de travaux également problématisent le rapport entre les éléments préexistants formalisés dans le référentiel d'évaluation et d'autres éléments informels qui interviennent dans le jugement de l'enseignant. Il en est de même pour ce qui concerne le rôle des contextes dans la définition de ces éléments et leur convocation parfois émergente dans les processus évaluatifs. En lien avec la thématique générale de ce numéro de la revue *Contextes et Didactiques*, l'objectif de cet article est de présenter quelques éléments de modélisation autour de ces questions à partir d'études empiriques préalables et de propositions de la littérature. Rappelons que la modélisation offre au travail scientifique une « construction abstraite pour comprendre le réel » (Figari, 1994 : 50), dont les unités d'analyse et les relations entre elles sont susceptibles d'être re-questionnées tout au long du processus de recherche en vue de leur confirmation, de leur ajustement ou de leur infirmation (Mottier Lopez et Figari, 2012).

Contrairement à un article de recherche classique qui commence par présenter le cadre théorique de l'étude, nous choisissons de renverser ici la logique, en faveur d'une approche *abductive* de la recherche (Anadón et Guillemette, 2007 ; Dechamboux et Mottier Lopez, 2015). À partir de données empiriques analysées dans une perspective exploratoire, nous réinterrogeons le modèle théorique pour mieux problématiser et comprendre comment les référents agissent dans les pratiques d'évaluation des enseignants et les contextes qu'ils convoquent. Des exemples concrets seront donnés, fondés sur des analyses dont les bases seront explicitées. Nous donnerons ainsi à voir les conceptions initiales à l'origine de ces analyses pour, ensuite, à la lumière des résultats obtenus, les remettre en partie en question. Ce faisant, cet article expose une démarche dans la recherche scientifique qui est souvent passée sous silence, à savoir la re-conceptualisation théorique du phénomène étudié à partir de données empiriques analysées.

Cet article commence par présenter notre recherche et son cadre d'analyse. Puis, à partir de l'exposition de constats principaux, nous réinterrogeons le modèle théorique initial autour de deux hypothèses de travail présentées successivement. Premièrement, nous posons l'hypothèse d'une relation dialectique entre référent et référé dans le jugement évaluatif en acte de l'enseignant¹. Rappelons que le référent est « ce par rapport à quoi un jugement de valeur est porté [...] le référent appartient à l'ordre des représentations d'objectifs » (Barbier, 1985 : 294). L'ensemble des référents forme le référentiel de l'évaluation (Figari, 1994). Quant au référé, il désigne « ce qui est constaté ou appréhendé de façon immédiate » (Lesne, 1984 : 132) par

¹« Jugement en acte » ou « activité évaluative » de l'enseignant. Dans cet article, les deux termes sont utilisés dans une relation de synonymie.

l'évaluateur ; il « désigne la partie de la réalité choisie » par ce dernier (Figari, 1994 : 44). À la suite des travaux scientifiques qui ont montré que des référents autres que ceux qui sont strictement spécifiques à l'objet évalué interviennent également dans les jugements évaluatifs des enseignants (par exemple, Allal et Mottier Lopez, 2008 ; Laveault, 2008), nous faisons l'hypothèse, deuxièmement, que la relation dialectique référent-référé dans le jugement de l'enseignant convoque des contextes pluriels dans les processus d'interprétation et de prise de décision évaluatives. Nous terminons en présentant la schématisation de notre modélisation renouvelée. Puis notre conclusion revient sur les points principaux de re-conceptualisation suggérée par notre modélisation et formule quelques axes de questionnement adressés à des travaux futurs autour de l'enjeu scientifique d'étudier la façon dont les différents éléments explicites et implicites du référentiel se combinent dans le jugement évaluatif de l'enseignant.

2. Une recherche sur les jugements évaluatifs des enseignants pendant qu'ils corrigent des travaux en français et en mathématiques

La recherche que nous avons entreprise² étudie, d'une part, les jugements évaluatifs des enseignants quand ils attribuent des notes certificatives aux travaux de leurs élèves en français (production écrite) et en mathématiques (résolution de problèmes) et, d'autre part, des pratiques dites de « modération sociale » entre ces mêmes enseignants visant à construire une culture d'évaluation commune. Dans cet article, nous ne nous intéresserons qu'au premier pan de la recherche, à savoir les jugements évaluatifs en acte des enseignants. Le détail du dispositif de recherche est présenté dans Mottier Lopez, Tessaro, Dechamboux et Morales Villabona (2012) ; nous renvoyons le lecteur intéressé par les pratiques de modération sociale à l'article de Tessaro, Dechamboux, Morales Villabona et Mottier Lopez (2016).

2.1. Participants et observation de leurs pratiques de correction

La recherche a impliqué deux cohortes d'enseignants de l'école primaire genevoise : 18 enseignants des degrés 5-6 HarmoS³ (élèves de 8 à 10 ans), et 13 enseignants des degrés 7-8 HarmoS (élèves de 10 à 12 ans). Conformément à ce qui avait été défini dans le contrat de recherche, chaque enseignant nous a contactés à quatre reprises pour que nous puissions assister à la correction de contrôles écrits de son cru, qu'il avait fait passer dans sa propre classe. Les travaux des élèves ne devaient pas être préalablement lus ou corrigés. Le règlement de l'enseignement primaire demande aux enseignants d'attribuer une note sur une échelle de un à six⁴. Sur chaque épreuve, les prescriptions institutionnelles exigent que l'enseignant insère un « entête » (ou champ de l'épreuve) pour indiquer les objectifs et/ou critères d'évaluation et le barème utilisé. Les épreuves ont été élaborées par les enseignants selon leur manière habituelle de procéder portant sur des objectifs d'apprentissage communs à chaque cohorte. Pour la recherche,

² Cette recherche est soutenue par le Fonds National Suisse de la recherche (n°100013-143453/1). Nous le remercions, ainsi que les enseignants qui ont participé et les personnes qui ont contribué à la recherche en plus des deux auteurs de cet article, Fernando Morales Villabona, Sophie Serry et Walther Tessaro.

³ Ce terme fait référence au concordat suisse sur l'harmonisation de la scolarité obligatoire entre les différents cantons suisses et accompagne la nouvelle désignation des degrés de la scolarité dans les cantons ayant approuvé ce texte. Disponible en ligne sur : <http://www.edk.ch/dyn/11737.php>.

⁴ Depuis la rentrée scolaire de 2014, les enseignants du cycle moyen de l'école genevoise doivent attribuer des notes de 1,5 à 6,0 par tranche de demi-notes et non plus par notes entières comme c'était le cas au moment de notre recherche.

une centration a été faite sur le processus évaluatif de chacun de ces enseignants tant au niveau des instruments utilisés (contrôles écrits proposant des tâches complexes) que de leur activité évaluative pendant la correction de ces mêmes épreuves.

Chaque observation⁵ s'est déroulée en trois temps successifs :

1. Un entretien semi-dirigé *ante*, fondé sur un canevas portant sur la description des objectifs visés, des choix évaluatifs, de l'inscription de l'évaluation dans le temps d'enseignement, ainsi que la description de chaque élève concerné.
2. Une verbalisation de l'enseignant pendant qu'il corrigeait quatre épreuves écrites d'élèves successivement, à partir de la méthode du *think-aloud* dans une approche qualitative (Charters, 2003). Les élèves choisis par l'enseignant devaient être d'un niveau scolaire moyen à faible dans la discipline concernée. Lafortune et Allal (2008) ont montré en effet que les enseignants tendent à davantage verbaliser avec ce profil d'élèves, l'évaluation étant moins « évidente ».
3. Un entretien semi-dirigé de clôture, sollicitant des remarques de l'enseignant à propos de la correction effectuée et des résultats obtenus par les élèves.

Les trois temps de l'observation ont été intégralement transcrits. Comme indiqué plus haut, ces transcriptions, mises en regard des traces écrites produites par l'enseignant pour son évaluation (par exemple, le corrigé de l'épreuve) et par l'élève (les épreuves elles-mêmes), ont fait l'objet de premières études exploratoires, notamment en examinant les critères d'évaluation formulés par l'enseignant dans l'entête de son épreuve comparativement aux propositions des moyens didactiques utilisés (plus particulièrement en français, Mottier Lopez, Serry et Sales Cordeiro, 2016) et la façon dont ces critères, ou d'autres non formalisés, sont mobilisés dans le jugement en acte de l'enseignant (également en français, Mottier Lopez, 2017 ; Mottier Lopez, Tessaro, Dechamboux, Morales Villabona et Serry, 2015). Les études se sont poursuivies en mathématiques également (Mottier Lopez et Dechamboux, 2016 ; Mottier Lopez et Morales Villabona, 2016). Nous présentons donc dans cet article un prolongement de ces études en nous préoccupant plus particulièrement de l'activité évaluative des enseignants au cours de la correction de ces contrôles écrits. Notre intention est de montrer comment les données analysées nous amènent à actualiser notre première modélisation.

2.2. Analyse du jugement en acte de l'enseignant et premiers constats

On pourrait penser aujourd'hui qu'il n'est peut-être plus nécessaire de démontrer que l'évaluation des apprentissages des élèves par un enseignant s'éloigne radicalement d'une opération de mesure. Comme le développe Hadji (1997) « l'évaluation est un acte de confrontation entre une situation réelle et des attentes concernant cette situation » (Hadji, 1997 : 38) ; elle « consiste à mettre en relation un référé et un référent » (Hadji, 1997 : 43). C'est à partir de cette conception, de plus en plus partagée par la communauté scientifique, que nous avons analysé nos données, en établissant une distinction, à la suite de Figari et Remaud (2014), entre :

- Le *référentiel explicite préexistant*, sous forme d'objectifs d'apprentissage, critères d'évaluation, indicateurs⁶ « déjà là ». Dans ce cas, les attentes sont formelles, communicables, pouvant être communes et partagées.

⁵ A quatre reprises, deux fois en français et deux fois en mathématiques, voir Mottier Lopez, Tessaro, Dechamboux et Morales Villabona (2012).

- Le *référentiel implicite*, qui émerge lors de la verbalisation de l'enseignant pendant la correction. Par exemple, l'enseignant admet que l'élève n'explique pas toutes les phases de son raisonnement mathématique dans une résolution de problèmes si celui-ci est correct. Dans ce cas, les attentes sont informelles, susceptibles d'être différentes entre les classes, voire entre les élèves d'une même classe.

Fortes de ces catégories, les premières analyses effectuées sur notre corpus visaient à mieux comprendre la relation entre les attentes explicites et implicites du référentiel, et à explorer la manière dont elles se combinent dans le jugement évaluatif de l'enseignant. Ce faisant, il s'agissait d'appréhender le processus de référentialisation qui implique, selon Figari (1994), une « recherche de références pertinentes (c'est-à-dire à la fois universelles et particulières) » (Figari, 1994 : 48) effectuée de manière consciente par l'évaluateur, susceptible de remettre en question des critères préexistants, de définir de nouveaux critères et, toujours, de spécifier les attentes. Car comme le soulignait déjà Figari dans ses travaux de 1994, le référentiel de l'évaluateur n'est pas fixé une fois pour toutes mais il est un *construit* susceptible de se reconstruire et de se transformer au fil de l'évaluation. Pour notre part, nous ajoutons « en acte » car nous cherchons à observer plus spécialement ce processus *au cours de l'activité évaluative de l'enseignant*. Nous élargissons donc la référentialisation au jugement évaluatif « en train de se faire », au-delà de la phase consciente et méthodique qui précéderait le jugement évaluatif.

Pour ce faire, le verbatim de la verbalisation de l'enseignant en cours de correction a été catégorisé à partir des unités conceptuelles suivantes :

- Le *référent* de l'évaluation, c'est-à-dire le critère d'évaluation ou tout autre « repère » verbalisé par l'enseignant ou inféré par le chercheur à partir des verbalisations, en observant dans quelle mesure ce référent était préexistant ou non.
- Le *référé* de l'évaluation, à savoir les « signes parlants » (Hadji, 2012) sélectionnés par l'enseignant dans le texte de l'élève et énoncés par l'enseignant en cours de correction ; cette deuxième unité permet d'identifier très précisément les éléments de la « réponse » de l'élève retenus par l'enseignant pour l'évaluation.
- Les *interprétations et raisonnements* verbalisés par l'enseignant donnant à voir en partie la confrontation entre référent et référé et la façon dont l'enseignant construit du sens et de la valeur à la réponse de l'élève.
- Les *décisions évaluatives* prises par l'enseignant, en termes de points attribués à chaque critère, puis concernant la note finale.

C'est à partir de cette analyse que nous présentons, dans la suite de cet article, quelques éléments de modélisation de la « référentialisation en acte » opérant dans le jugement évaluatif de l'enseignant.

3. Eléments d'une modélisation de la « référentialisation en acte » dans le jugement évaluatif de l'enseignant, comme processus de co-constitution référentielle située

Cette partie commence par présenter quelques constats issus des analyses, à partir desquels nous exposons ensuite des éléments de notre modélisation fondée sur deux hypothèses de travail qui amènent à définir la *référentialisation en acte* comme étant non plus seulement une modification

⁶ Ces trois notions sont évidemment à distinguer entre elles.

et mise à jour du référentiel pendant l'évaluation, mais comme étant aussi un processus dynamique de co-constitution :

- entre référents et référés ;
- foncièrement située, c'est-à-dire par rapport, et avec, des contextes significatifs agissant.

La partie se termine par l'exposition schématique de la modélisation à laquelle nous aboutissons et qui, comme toute modélisation scientifique, a pour vocation d'évoluer et d'être questionnée par des nouveaux travaux scientifiques.

3.1. Constats autour de la distinction entre le référentiel de l'évaluation préexistant et les référents-référés agissant dans le jugement évaluatif de l'enseignant

Les constats issus de nos études⁷ donnent à voir la complexité du jugement évaluatif de l'enseignant, y compris quand celui-ci s'appuie sur des critères explicités dans l'entête de l'épreuve certificative, que ce soit en termes :

- De *critères de réalisation*, désignant les qualités des processus ou démarches à l'œuvre pour réaliser la production attendue (Bonniol et Genthon, 1989 ; Veslin et Velsin, 2001). Dans notre étude, cela concerne l'écriture de différents genres textuels en français en respectant diverses composantes propres à chaque genre et la résolution de problèmes additifs et multiplicatifs en suivant un certain nombre d'étapes, dans des épreuves certificatives réunissant plusieurs « exercices » ou tâches évaluatives.
- De *critères de réussite* portant sur le résultat (Veslin et Veslin, 2001), qui se traduisent dans les épreuves certificatives de nos enseignants par des points, des seuils de réussite par objectif ou par critère, par la fixation d'un seuil de réussite global matérialisant le niveau d'exigence attendu.

Dans leurs entêtes, les enseignants utilisent des formulations variées entre ce qui relève littéralement d'un objectif d'apprentissage tel qu'il est proposé dans le plan d'études officiel (*résoudre des problèmes additifs et multiplicatifs...*), d'objectifs qui sont reformulés par eux, parfois à la première personne (du point de vue de l'élève), souvent proches de la formulation d'un critère d'évaluation (*je fournis au minimum trois arguments*), parfois donnant à voir des indicateurs au sein même de la formulation (*j'organise mon texte au moyen d'organiseurs tels que : tout d'abord, par ailleurs, ensuite, pour terminer*). Les annexes 1 et 2 présentent respectivement un exemple d'entête en français, pour l'écriture d'un texte argumentatif (élèves de 9-10 ans), et un exemple d'entête en mathématiques (élèves de 9-10 ans), typiques des pratiques des enseignants de notre recherche. Pour des questions d'écriture, par la suite, nous parlerons des « critères d'évaluation » de l'entête, tout en attirant l'attention sur le fait que les formulations des enseignants combinent fréquemment des objectifs, critères et indicateurs en français, et privilégient plutôt la formulation d'objectifs en mathématiques avec des critères implicites qui apparaissent pendant la correction (par exemple : *raisonnement mathématique pertinent, résultat final correct*).

⁷ Le format de cet article ne permet pas d'exposer dans le détail l'ensemble des résultats des études que nous avons citées. Nous ne reprenons ici que les constats principaux.

Nos analyses montrent différents cas de figure dans la façon dont les enseignants s'appuient sur les critères formulés dans leurs entêtes pour évaluer les productions des élèves⁸. Ces cas montrent une grande diversité dans leur rapport avec le référentiel explicite préexistant. Nous avons tenté d'ordonner ci-dessous les principaux types de cas en fonction d'un rapport de moins en moins étroit avec les éléments préexistants du référentiel formalisé dans les entêtes :

- L'enseignant effectue une mise en relation rapide entre le critère d'évaluation et le référent retenu dans la production de l'élève ; le ou les points sont rapidement attribués sans verbalisations particulières.
- L'enseignant se pose des questions par rapport à ses attentes liées à un critère au regard d'un élément de la production de l'élève (un référent) inattendu, singulier, partiellement correct, difficile à comprendre. L'enseignant reformule alors le critère, précise le sens de celui-ci par rapport à la production particulière de l'élève, parfois aussi par rapport aux autres productions d'élèves déjà corrigées, ou aux prochaines à venir.
- Deux critères distincts sont formulés (soit dans l'entête, soit dans la verbalisation) mais, dans la confrontation à la production réelle de l'élève, l'enseignant peine à les distinguer ; les critères apparaissent finalement liés les uns aux autres dans certaines productions d'élèves. À nouveau, l'enseignant (se) précise (à lui-même) ses attentes propres à chaque critère et ce qu'il interprète de la production de l'élève notamment quand celle-ci garde une certaine opacité (par exemple, un raisonnement mathématique partiellement correct ; une structuration textuelle qui influence le contenu thématique du texte rédigé ou inversement).
- L'enseignant a formulé dans le même critère deux dimensions et envisageait préalablement d'attribuer un point par dimension. Par exemple : *mon histoire est divertissante, il y a de la fantaisie*. Au cours de son jugement en acte, l'enseignant n'évalue finalement qu'une seule des dimensions énoncées ou choisit de les regrouper. Il est dans l'impossibilité de les différencier avec des indicateurs précis ou de lever un implicite dans la production de l'élève qui lui permettrait de le faire.
- Pendant la correction, d'autres référents se mêlent aux critères explicités dans l'entête et prennent finalement le dessus, tels *le progrès de l'élève entre le brouillon initial du texte à rédiger et le texte final* ou *l'effort de l'élève*, pour attribuer les points à un critère formel portant, par exemple, sur le *respect de l'orthographe*.
- Au regard d'une production originale ou non attendue d'un élève ou de plusieurs élèves, l'enseignant convoque un référent non anticipé dans son entête, par exemple en mathématiques afin de faire une différence entre une *petite erreur de calcul* et un *raisonnement partiellement correct* ou encore quand *les calculs sont corrects dans un raisonnement incorrect*.
- Enfin, citons le cas de figure dans lequel c'est l'atteinte du seuil de réussite qui apparaît comme déterminante⁹ et qui amène l'enseignant à revoir certaines de ses décisions intermédiaires à propos de critères pour lesquels il avait hésité à attribuer ou non des points. Le référent pour l'enseignant est ici de considérer que, si ses hésitations ont une conséquence négative sur la note finale, l'hésitation doit alors être *au bénéfice de l'élève*, notamment quand il considère que l'élève *mérite* d'avoir la note minimale. Le mérite, en tant que référent émergent, oriente alors la décision évaluative de l'enseignant.

⁸ Dans cet article, nous n'aurons pas la place de discuter les différences entre le français et les mathématiques.

⁹ Note finale de l'élève en-dessous ou non de la note de réussite minimale de 4. Voir annexes.

Parmi ces différents cas de figure, non exhaustifs, nous retiendrons que la dynamique référentielle est complexe entre :

- des critères préexistants qui se spécifient, voire se modifient, au contact des productions d'élèves à évaluer ;
- des référents informels préexistants qui se combinent à des critères formalisés (par exemple, la prise en compte des efforts d'un élève, ou de son « mérite ») ;
- des référents qui émergent, inattendus, afin de pouvoir faire face à la singularité d'une ou plusieurs productions d'élève ;
- des critères qui apparaissent tout particulièrement déterminants (notamment l'atteinte du seuil de réussite) et qui influencent l'appréciation liée à d'autres critères d'évaluation.

Par rapport à notre projet de modélisation conceptuelle, notamment relativement aux référents dits préexistants, ces constats nous amènent désormais à ne parler de référents que lorsque ceux-ci sont *effectivement* mobilisés dans le jugement évaluatif en acte de l'enseignant, c'est-à-dire quand ils deviennent *ce par rapport à quoi* la production de l'élève est appréciée et qui orientent la prise de décision évaluative (Mottier Lopez, 2017). Ce choix vise à différencier ce qui relève d'attentes et de repères préexistants (sous forme d'objectifs d'apprentissage, critères d'évaluation, indicateurs) mais qui, dans le jugement évaluatif, sont susceptibles de ne pas être réellement convoqués, ou dont le sens peut être transformé, partiellement ignoré, combiné, etc.

Nos analyses montrent également que le sens des référents, qui font *système* entre eux comme on l'a vu plus haut, se construit au contact des productions réelles des élèves, plus spécifiquement par rapport aux indices ou « signes parlants » (Hadj, 2012) que l'enseignant sélectionne dans la production de l'élève à des fins d'interprétation et de prise de décision évaluatives. Les pratiques de modération sociale que nous avons observées (Tessaro, Dechamboux, Morales Villabona et Mottier Lopez, 2016), qui consistent à demander aux enseignants de confronter leurs évaluations pour construire des consensus sur la note à attribuer à l'élève (Wyatt-Smith, Klenowski et Gunn, 2010), montrent que les enseignants¹⁰ ne donnent pas sens aux mêmes éléments de la production de l'élève, notamment quand celle-ci est complexe. Autrement dit, le référé relève aussi d'une construction dans le jugement interprétatif de l'enseignant. Ceci nous amène à réserver également le concept de référé à l'unité que l'enseignant retient et interprète effectivement de la production de l'élève qu'il est train d'évaluer. Nous utilisons le concept d'indicateur pour désigner les éléments formulés et anticipés par l'enseignant avant sa correction, en termes de repères concrets à observer dans les productions des élèves.

La figure 1 synthétise ces premiers choix conceptuels, puis nous poursuivons sur nos hypothèses d'une relation dialectique entre référés et référents en lien avec des contextes agissant.

¹⁰ Il s'agit des mêmes enseignants que ceux de notre étude sur les jugements évaluatifs en actes, à propos des mêmes entêtes (telles celles présentées dans les annexes) et épreuves certificatives.

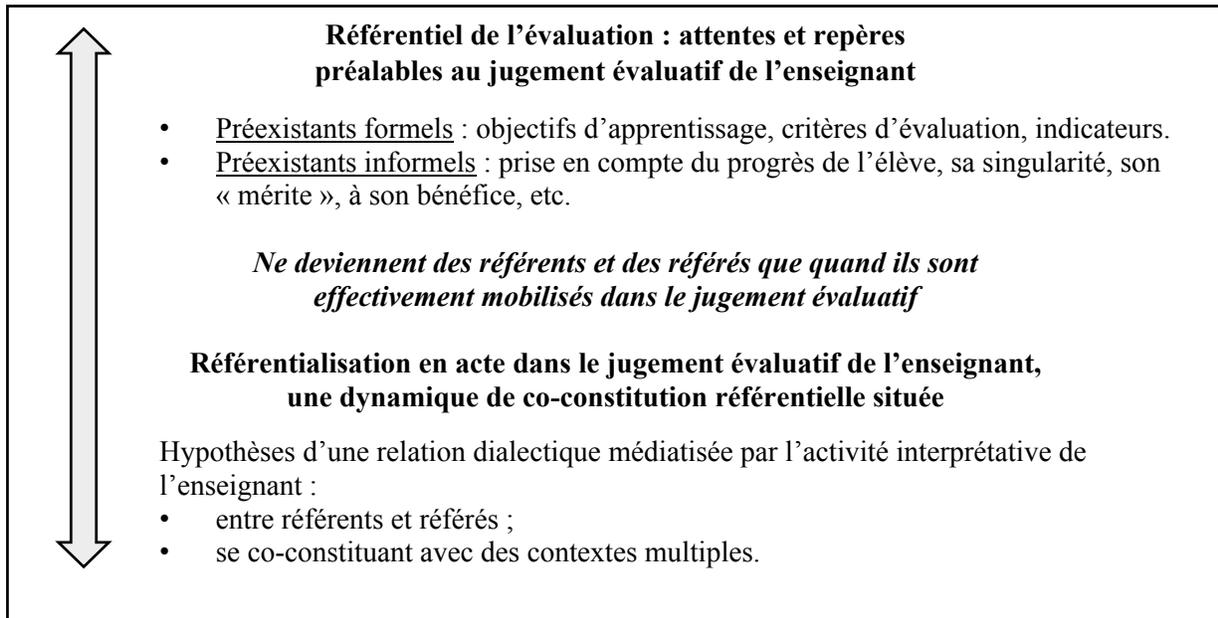


Figure 1 : Premiers éléments conceptuels fondant notre proposition de modélisation

3.2. Conceptualisation d'une relation dialectique entre référent et référé...

Dans la littérature scientifique, la relation entre référent et référé a été différemment conceptualisée, notamment en termes de « mesure d'écart » dans une approche d'évaluation-mesure, ou en termes de « comparaison », voire de « confrontation » dans une approche qualitative d'évaluation consistant à déterminer une valeur¹¹ (voir par exemple Goasdoué et Vantourout, 2017). Quelle que soit la conceptualisation, chaque terme est ici appréhendé comme une unité distincte de l'autre, dont le sens lui reste propre.

Notre hypothèse de travail est plutôt celle d'une *relation dialectique* (au sens de Lave, 1988) entre référent et référé, c'est-à-dire vus comme étant co-constitutifs l'un de l'autre¹². Concrètement cela signifie que le sens des référents et référés dans le jugement en acte de l'enseignant s'établit dans leur mise en rapport et constitutions réciproques. Il s'agit d'une relation *indirecte*, comme montré dans les travaux de Lave et successeurs (Cobb et Bowers, 1999 ; Greeno, 1997 ; Lave, 1988 ; Lave et Wenger, 1991 ; Mottier Lopez, 2008) : la relation dialectique suppose en effet des *processus demédiation* qui, dans notre étude, relèvent de l'activité interprétative et de prise de décision de l'enseignant.

Cette conceptualisation amène à considérer que l'évaluation de productions complexes s'éloigne de l'idée de rechercher des « correspondances » entre des attendus et des réponses d'élèves, comme montré également par d'autres travaux (par exemple, Vial, 2012). Dans notre recherche, les attentes préexistantes du référentiel de l'évaluation (objectifs, critères, indicateurs formulés dans les entêtes des épreuves) fournissent des catégories importantes qui structurent le jugement

¹¹ Détermination d'une valeur qui n'exclut pas l'usage de nombres (des points, des notes chiffrées) mais qui se fonde sur un raisonnement qualitatif et interprétatif qui n'est pas référencé à un modèle quantitatif de la mesure.

¹² Mais sans pour autant verser dans un syncrétisme réducteur qui amènerait à ne plus les différencier.

évaluatif de l'enseignant. Ces catégories conservent cependant une forme de variation ou d'instabilité irréductible. Contrairement à la conception d'une évaluation-mesure, cette instabilité n'est pas *de facto* considérée comme un biais ou une limite du jugement évaluatif de l'enseignant. Dans une approche située de l'évaluation, telle celle que nous développons dans nos travaux, cette instabilité est vue comme intrinsèque à la relation dialectique qui se construit entre référents et référés, et si besoin se déconstruit / se reconstruit, s'adapte à la singularité des productions des élèves. C'est dans ce processus dynamique d'interprétation et de construction de sens et de valeurs propre à l'évaluation (Vial, 2012) que l'enseignant tente de stabiliser ses référents et référés, fondés sur des éléments contextuels préexistants (*l'arena*, dans la terminologie de Lave, 1988, qui relèverait par exemple des prescriptions des plans d'études, des formulations dans les entêtes des épreuves) et émergents (*le setting*, au sens de Lave). Cette recherche de stabilisation, cette mise en relation, qui peut même s'avérer parfois créatrice, peut alors amener l'enseignant à vouloir modifier son référentiel d'évaluation dont certains éléments lui semblent inadaptés, soit de façon explicite (nouvelle formulation d'un critère ou modification du nombre de points dans l'entête de l'épreuve ou pour la prochaine épreuve ; décidés pour tous les élèves ou pour certains, comme ceux à besoins éducatifs particuliers par exemple), soit de façon implicite (c'est-à-dire en spécifiant ou transformant le sens d'un critère mais sans forcément le rédiger et le communiquer ; soit pour tous les élèves ou seulement pour certains élèves ayant un profil particulier).

3.3. ... au cœur d'un jugement évaluatif foncièrement situé

La dernière hypothèse examinée dans cet article découle directement de la conception de la *relation dialectique* développée dans une perspective de l'apprentissage situé (par exemple, Cobb et Bowers, 1999 ; Greeno, 1997 ; Lave, 1988 ; Lave et Wenger, 1991 ; Mottier Lopez, 2008). Cette perspective postule en effet un caractère foncièrement incarné aux activités sociales et individuelles (le jugement évaluatif de l'enseignant, dans notre recherche), situées dans des contextes socioculturels, dans *le hic et nunc* de la situation expérimentée (ici, la situation *immédiate* de correction des productions des élèves et ses dimensions contextuelles) et également par la convocation mentale de contextes *non immédiats* mais agissant néanmoins sur l'activité humaine (par exemple, le contexte de la classe, celui de l'établissement scolaire, le contexte familial d'un élève, etc.). L'hypothèse, qui fonde notre proposition de modélisation, est que les significations produites par le moyen du jugement évaluatif de l'enseignant sont également le produit des situations expérimentées et de leurs contextes. Réciproquement, ces situations et leurs contextes sont elles-mêmes constituées et susceptibles d'être transformées en partie par l'activité évaluative motivée par des buts, des intentions, des pratiques sociales et culturelles (Mottier Lopez, 2016).

Nos analyses fines des verbalisations des enseignants, croisées aux traces écrites produites par les élèves et par eux-mêmes, montrent que la construction de significations dans la relation dialectique référent-référé est résolument liée à des contextes qui peuvent être pluriels et qui participent à part entière à la dynamique de co-constitution référentielle. Prenons un exemple emblématique de cette co-constitution : en mathématiques, une enseignante accepte une réponse d'élève non canonique (sans formulation mathématique, mais dont le raisonnement est correct et qui débouche sur une réponse correcte). L'enseignante justifie : *donc moi je compte juste, parce que je leur ai dit qu'ils faisaient comme ils voulaient* (consignes orales transmises aux élèves dans le contexte de passation de l'épreuve et/ou contrat didactique plus général, la procédure experte n'a pas encore été institutionnalisée dans le contexte de la classe). Pour le même

problème, un autre élève de la même classe a tenté, quant à lui, une résolution mathématique en posant des algorithmes pertinents mais non maîtrisés. La réponse annoncée est incorrecte. L'élève *perd un point*. L'implicite, lié au contrat didactique et normes de la microculture de classe (Mottier Lopez, 2008), est ici qu'il vaut mieux pour les élèves proposer un raisonnement que l'on maîtrise, quitte à ne pas mobiliser certains outils mathématiques, plutôt que de proposer des procédures plus expertes mais non maîtrisées – tout au moins à ce moment précis de la progression didactique de la classe.

Prenons un autre exemple significatif en français, la co-constitution référentielle s'observant dans les deux disciplines concernées par notre recherche¹³. Au cours de sa verbalisation, une autre enseignante convoque le critère-référent *pour éviter les répétitions et les ambiguïtés on désigne les personnages de manière claire* et elle pointe plusieurs référés dans le texte de l'élève : *L'enfant de 15 ans [...] le gamin*. Dans le texte de l'élève, les reprises anaphoriques sont peu nombreuses et l'anaphore « le gamin » apparaît peu adéquate aux yeux de l'enseignante quant aux normes langagières scolaires. Cependant, l'enseignante interprète sa mise en relation référent-référent au regard du contexte de la classe et du contrat didactique négocié avec ses élèves : *on avait insisté sur le fait qu'on ne voulait pas tout le temps 'un enfant de 15 ans' [...] le but c'était qu'ils changent*. Cette interprétation se fait dans une relation de couplage avec l'activité manifestée par l'élève en rédigeant son texte : *il a fait un effort même si ce n'est pas top[...] il a fait un effort de changer*. Cet exemple montre le caractère situé de la dynamique référentielle qui sous-tend le jugement évaluatif de l'enseignant, ici pour ce qui concerne plus spécialement l'interprétation de la production de l'élève sur le plan de son usage des reprises anaphoriques :

- *pour éviter les répétitions et les ambiguïtés on désigne les personnages de manière claire* (dans l'entête du contrôle écrit, critère préexistant lié au modèle de l'objet de savoir en jeu, ici le genre de texte) ;
- *on ne voulait pas tout le temps un enfant de 15 ans* (contrat didactique négocié entre l'enseignant et ses élèves dans le contexte de la classe) ;
- *effort de l'élève* (implicite valorisé par l'enseignant relatif à l'implication active de l'élève).

Dans son entête, l'enseignante avait décidé d'attribuer un seul point au critère d'évaluation concerné. Dans son jugement en acte, elle souhaiterait disposer de deux points, ce qui lui permettrait de valoriser « l'effort » de l'élève en lui octroyant un point sur deux. Si l'attente relative à ce critère n'est pas totalement remplie par l'élève, sa production ne « mérite » cependant pas de n'obtenir aucun point. Mais cette décision évaluative exige, selon l'enseignante, une concertation avec la collègue avec qui elle élabore les épreuves certificatives : *j'appelle ma collègue après pour voir si elle me suit là-dedans ou pas*. Cette décision semble acceptable car *la demande de la directrice* (de l'établissement scolaire), *c'est implicite, mais c'est de pas couler les élèves*. Ainsi, la prise de décision dans le jugement évaluatif de l'enseignant est, elle également, liée à des contextes, ici de collaboration professionnelle entre enseignants et de culture d'établissement scolaire.

¹³ Pour les lecteurs intéressés, ce cas est présenté dans son intégralité dans Mottier Lopez (2017).

En bref, nous retiendrons que les processus d'interprétation et de prise de décision qui sous-tendent le jugement en acte de l'enseignant s'inscrivent dans des contextes pluriels qui participent pleinement à la dynamique de co-constitution référentielle qui nous intéresse dans cet article.

3.4. Essai de modélisation de la référentialisation en acte, comme processus de co-constitution référentielle située

Au regard des développements proposés dans cet article, la figure 2 synthétise les principaux éléments de modélisation de la référentialisation en acte que nous proposons, en tant que processus dynamique de co-constitution référentielle située. Elle demandera, par la suite, à s'intégrer à un modèle plus général de l'activité évaluative située.

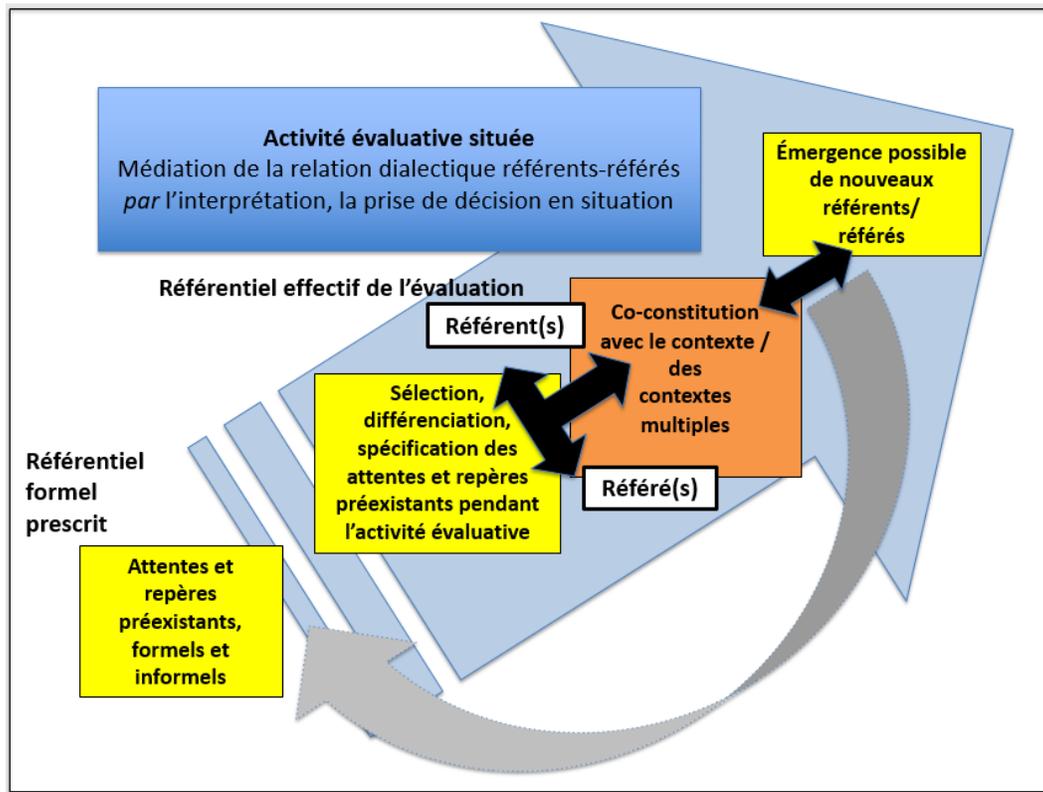


Figure 2 : Modélisation de la référentialisation au cours de l'activité évaluative

Notre modélisation conceptualise, en termes de système, une double relation dialectique (flèches noires dans la figure) entre référents et référés au cœur du jugement évaluatif, et avec les contextes possiblement pluriels de l'activité évaluative. Ce faisant, notre modélisation met radicalement en question une conception positiviste, encore courante dans la littérature contemporaine en évaluation bien que parfois diffuse, d'un rapport référent-référé qu'il suffirait de « comparer » pour, notamment, assurer l'objectivité et l'équité de l'évaluation des apprentissages des élèves. Elle argumente au contraire la nature co-construite des référents-référés et de leur mise en relation dans le jugement interprétatif de l'enseignant. Ce jugement, en tant qu'activité foncièrement située, intègre des éléments de sens propres aux contextes signifiants pour les interprétations et décisions évaluatives à prendre. Ce peut être le contexte immédiat à l'activité évaluative et sa situation, mais aussi des contextes distaux convoqués par

l'activité interprétative de l'évaluateur, liés à ses expériences et vécus en tant qu'enseignant, évaluateur, élève, etc., liés à une culture d'établissement, à un contexte familial de l'élève, etc. C'est en cela que nous proposons de parler de *multiréférentialité* du jugement évaluatif de l'enseignant. Dans des précédents travaux, nous avons déjà convoqué ce concept, notamment dans le cadre du jugement professionnel de l'enseignant en évaluation (par exemple, Mottier Lopez et Allal, 2008), mais sans l'observer aussi finement dans les processus de jugement en acte et dans le rapport constitutif avec les référents-référés et contextes convoqués.

Notre modélisation pointe également la dimension processuelle de cette mise en relation, en prenant en considération la façon dont les attentes préexistantes formelles (référentiel prescrit de l'évaluation) et informelles sont effectivement prises en compte dans le jugement évaluatif de l'enseignant (sélectionnées, parfois écartées, différenciées en fonction des productions des élèves et/ou de leurs situations possiblement singulières) et toujours spécifiées comme l'argumentent Figari et Remaud (2014). La thèse de la relation d'indissociabilité entre activités et contextes amène ici à considérer que tous référents, y compris ceux qui sont préexistants et formels, relèvent « toujours en partie d'une interprétation située qui les spécifie tout en devant garantir un niveau de généralisation qui puisse concerner l'ensemble des élèves (d'une même classe, et au-delà quand les enseignants se coordonnent entre eux) » (Mottier Lopez, 2017 : 187). Ce faisant, « il est légitime de parler aussi de référents émergents à leur propos » (Mottier Lopez, 2017 : 187). Un des éléments de notre modélisation est aussi de considérer l'émergence de nouveaux référents-référés au cours du jugement évaluatif et la façon dont ils opèrent pendant l'activité en train de se faire et dans des situations d'évaluation ultérieures – en tant qu'attentes pouvant devenir formelles ou restant dans un informel conscient par exemple – comme la prochaine production d'élève à corriger ou pour la prochaine situation d'évaluation comparable.

4. Discussion conclusive

Depuis les travaux de Barbier (1985), Figari (1994), Hadji (1997) et Lesne (1984) notamment, les concepts de référent et de référé sont régulièrement invoqués dans les travaux scientifiques en évaluation, mais sans être pleinement stabilisés par rapport aux concepts d'attente, objectif, critère, indicateur, y compris entre ce qui relève du prescrit, préexistant, « représenté », et ce qui relève de l'activité évaluative en train de se faire, du concret, du singulier, de l'agissant. A partir de cette littérature existante et de nos études exploratoires, nous avons choisi, dans cet article, de soumettre à la communauté scientifique une modélisation qui explicite la relation que nous établissons entre ces différentes unités conceptuelles. Cette modélisation demandera à se développer, s'affiner, se préciser, et évidemment à être mise en question – le propre du travail scientifique – avec des nouvelles recherches et débats qu'elle suscitera certainement.

L'analyse de nos données empiriques nous a amenés à observer que non seulement certaines des attentes explicitées dans le référentiel d'évaluation ne sont pas toujours « ce par rapport à quoi » le jugement de l'enseignant opère de manière effective, mais que d'autres attentes plus ou moins implicites peuvent également émerger au cours du jugement en acte au regard de référés inattendus par exemple. C'est pourquoi, notre modélisation propose de parler de référents et référés que lorsque ceux-ci sont *effectivement* mobilisés dans l'activité évaluative de l'enseignant. Cette distinction nous paraît particulièrement importante pour mieux appréhender les phénomènes de référentialisation qui se jouent dans la complexité de l'activité évaluative. Elle amène à poser certaines questions à propos, entre autres :

- Des attentes affichées dans un référentiel d'évaluation mais qui s'avèrent non opérantes dans le jugement évaluatif de l'enseignant. À quoi servent-elles ? Pourquoi le sont-elles ? Servent-elles, par exemple, juste à répondre à des prescriptions institutionnelles ? Ou est-ce plutôt lié à un manque de connaissances liées à l'objet de savoir en jeu dans l'évaluation et des critères qui peuvent être réellement évalués dans une épreuve ? Etc.
- Des attentes implicites mais néanmoins préexistantes et celles affichées dans l'entête. Pourquoi certaines ne peuvent-elles pas être formalisées dans le référentiel de l'évaluation, sachant aussi que tout ne peut pas être explicité (Hameline, 1979) ? Dans quelle mesure est-ce fonctionnel pour l'enseignant de garder une certaine souplesse, marge de manœuvre ? Dans quelle mesure certaines de ces attentes sont-elles indicibles, en quoi, pourquoi ? Etc.
- Des référents émergents au cours du jugement de l'enseignante. Qu'est-ce qui les fait émerger ? Dans quelle mesure l'enseignant choisit-il, ou non, de les « sortir » de l'implicite pour les faire figurer dans son référentiel d'évaluation par exemple ? À quelles conditions ? Pourquoi certains référents émergents ne peuvent-ils pas, aux yeux de l'enseignant, se formaliser dans le référentiel de l'évaluation ? Etc.

Un risque majeur d'une lecture rapide de notre conceptualisation serait de conclure que l'évaluation serait toujours relative, instable, mouvante, subjective. Nos analyses montrent que la recherche de stabilité fait aussi pleinement partie de la dynamique de co-constitution entre référents-référés et contextes agissants dont nous avons vu qu'ils étaient multiples entre interprétations et prises de décision évaluatives. Dans cette perspective, l'objectivité et l'équité de l'évaluation demandent à être également considérées comme des construits – ou des produits – de la dynamique de co-constitution référentielle. À nos yeux, un enjeu scientifique est alors de mieux appréhender les conditions de cette construction, nécessitant de toujours mieux comprendre en situation d'évaluation les systèmes référents-référés mobilisés et les contextes pluriels convoqués par elle.

Enfin, pour conclure, il est important de préciser que certains des phénomènes que nous avons pointés sont déjà bien documentés dans la littérature scientifique, tels que la prise en compte de l'effort de l'élève et de son mérite, le choix de l'enseignant de ne pas attribuer une note globale en dessous du seuil de réussite quand celle-ci ne correspond pas à sa représentation de l'élève, etc. Malgré leur pertinence, nous avons cependant choisi de ne pas citer ces travaux, souvent d'orientations psychologique ou sociologique. Dans le cadre restreint de cet article, notre objectif était en effet de rester au plus près des catégories conceptuelles propres à une *théorie de l'évaluation* (Figari et Remaud, 2014 ; Mottier Lopez et Figari, 2012), plus spécialement du point de vue des phénomènes de référentialisation en acte. Des mises en perspective entre les travaux seront évidemment à envisager pour une appréhension plus large des questions que nous avons soulevées.

Références bibliographiques

- Allal, L. et Mottier Lopez, L. (2005). Formative Assessment of Learning : A Review of Publications in French. Dans *Formative Assessment – Improving Learning in Secondary Classrooms* (p. 241-264). Paris : OECD Publication.
- Anadón, M. et Guillemette, F. (2007). La recherche qualitative est-elle nécessairement inductive ? *Recherches Qualitatives*, 5, 26-37.

- Bain, D. et Schneuwly, B. (1993). Pour une évaluation formative intégrée dans la pédagogie du français : de la nécessité et de l'utilité de modèles de référence. Dans L. Allal, D. Bain et P. Perrenoud (dir.), *Évaluation formative et didactique du français* (p. 51-79). Neuchâtel : Delachaux et Niestlé.
- Barbier, J.-M. (1985). *L'évaluation en formation*. Paris : Presses Universitaires de France.
- Bonniol, J.-J. et Genthon, M. (1989). L'évaluation et ses critères : les critères de réalisation. *Repères pour la rénovation de l'enseignement du français*, 79, 106-115.
- Charters, E. (2003). The use of think-aloud methods in qualitative research. An introduction to think-aloud methods. *Brock Education*, 12, 68-82.
- Cobb, P. et Bowers, J. (1999). Cognitive and situated learning perspectives in theory and practice. *Educational Researcher*, 3, 4-15.
- Cros, F. et Raisky, C. (2010). Référentiel. *Recherche et formation*. Mis en ligne le 01 mai 2010, consulté le 13 février 2017. URL : <http://rechercheformation.revues.org/215>
- Dechamboux, L. et Mottier Lopez, L. (2015). L'enquête de l'enseignant-évaluateur et du chercheur : l'abduction au cœur du processus et quelques raisons de les comparer. *Recherches Qualitatives*, 34 (1), 173-195.
- Figari, G. (1994). *Évaluer : quel référentiel ?* Bruxelles : De Boeck.
- Figari, G. et Remaud, D. (2014). *Méthodologie d'évaluation en éducation et formation*. Bruxelles : De Boeck.
- Gagnebin, A., Guignard, N. et Jaquet, F. (1997). *Apprentissage et enseignement des mathématiques, commentaires didactiques sur les moyens d'enseignement pour les degrés 1 à 4 de l'école primaire*. Neuchâtel : Institut de recherche et de documentation pédagogique.
- Gouasdoué, R. et Vantourout, M. (2016). Evaluations scolaires et étude du jugement des enseignants : pour une docimologie cognitive. Dans P. Detroz, M. Crahay et A. Fagnant (dir.). *L'évaluation à la lumière des contextes et des disciplines*. (p. 141-168). Bruxelles : De Boeck.
- Greeno, J. G. (1997). On claims that answer the wrong questions. *Educational Researcher*, 26(1), 5-17.
- Hadji, C. (1997). *L'évaluation démystifiée : mettre l'évaluation scolaire au service des apprentissages*. Paris : ESF.
- Hadji, C. (2012). *Faut-il avoir peur de l'évaluation ?* Bruxelles : De Boeck.
- Hameline, D. (1979). *Les objectifs pédagogiques en formation initiale et continue*. Paris : ESF éditions.
- Lafortune, L. et Allal, L. (dir.). (2008). *Jugement professionnel en évaluation : pratiques enseignantes au Québec et à Genève*. Québec : Presses de l'Université du Québec.
- Lave, J. (1988). *Cognition in practice : Mind, mathematics and culture in everyday life*. Cambridge, UK : Cambridge University Press.
- Lave, J. et Wenger, E. (1991). *Situated learning : Legitimate peripheral participation*. Cambridge, MA : Cambridge University Press.
- 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.
- Lesne, M. (1984). *Lire les pratiques de formation d'adultes – essai de construction théorique à l'usage des formateurs*. Paris : EDILIG.
- Mottier Lopez, L. (2008). *Apprentissage situé : la microculture de classe en mathématiques*. Berne : Peter Lang.
- Mottier Lopez, L. (2015). *Evaluations formative et certificative des apprentissages. Enjeux pour l'enseignement*. Bruxelles : De Boeck.

- Mottier Lopez, L. (2016). La compétence à l'école pensée à partir de la perspective située de l'apprentissage. *Education et francophonie*, 44(2), 152-171. Disponible en ligne sur : http://www.acelf.ca/c/revue/pdf/EF-44-2-152_MOTTIER_LOPEZ.pdf
- Mottier Lopez, L. (2017). *Une modélisation pour appréhender la référentialisation dans l'évaluation des apprentissages des élèves*. Dans M. Crahay, A. Fagnant et P. Detroz (dir.), *L'évaluation à la lumière des contextes et des disciplines* (p. 169-192). Bruxelles : De Boeck.
- Mottier Lopez, L. et Allal, L. (2008). Le jugement professionnel en évaluation : un acte cognitif et une pratique sociale située. *Revue suisse des sciences de l'éducation*, 3, 465-482.
- Mottier Lopez L. et Dechamboux, L. (2016). *La place et le rôle de la référence dans les jugements évaluatifs des enseignants*. Communication au congrès annuel 2016 de la SSRE, Université de Lausanne, Suisse.
- Mottier Lopez, L. et Figari, G. (2012). Modèles et modélisations face à une crise du sens de l'évaluation en éducation. Dans L. Mottier Lopez et G. Figari (dir.), *Modélisations de l'évaluation en éducation : questionnements épistémologiques* (p. 7-23). Bruxelles : De Boeck.
- Mottier Lopez, L. et 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. et Morales Villabona, F. (2016). *Étude de la pratique évaluative de cinq enseignants quand ils corrigent des productions d'élèves en français et en mathématiques : entre hésitations, difficultés et jugement professionnel*. Communication au congrès annuel 2016 de la SSRE, Université de Lausanne, Suisse.
- Mottier Lopez, L., Serry, S. et Sales Cordeiro, G. (2016). Evaluation certificative de productions écrites d'élèves : référence au modèle didactique du genre textuel dans l'épreuve et lors de la correction de « faits divers ». *Recherche en didactique du français*, 8, 27-61.
- Mottier Lopez, L., Tessaro, W., Dechamboux, L. et 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). Disponible en ligne sur : <http://questionsvives.revues.org/1235>.
- Mottier Lopez, L., Tessaro, W., Dechamboux, L., Morales Villabona, F. et Serry, S. (2015, août). *Situated judgment in summative assessment practices about mathematics and writing in primary school*. Communication in the symposium organized by L. Mottier Lopez « From teachers' assessment practices to shared professional judgment cultures », EARLI, Chypre.
- Tessaro, W., Dechamboux, L., Morales Villabona, F. et Mottier Lopez, L. (2016). La modération sociale : pour se former à l'évaluation des apprentissages. *Education permanente*, 208(3), 65-76.
- Velsin, O. et Veslin, J. (2001). Evaluation formatrice et critères de réalisation. Dans G. Figari et M. Achouche (dir.), *L'activité évaluative réinterrogée : regards scolaires et socioprofessionnels* (p. 89-101). Bruxelles : De Boeck.
- Vial, M. (2012). *Se repérer dans les modèles de l'évaluation. Méthodes, dispositifs, outils*. Bruxelles : De Boeck.
- Wyatt-Smith, C., Klenowski, V. et Gunn, S. (2010). The centrality of teachers' judgement practice in assessment: a study of standards in moderation. *Assessment in Education : Principles, Policy et Practice*, 17(1), 59-75.

Annexe 1 : Exemple d'entête, écriture d'un texte argumentatif (élèves de 9-10 ans)

Points obtenus :	23	Seuil de réussite :	20/30	Note obtenue :	5
GUIDE DE PRODUCTION ET D'EVALUATION					
	Objectifs	Elève :	contrôlé	Enseignant :	contrôlé
Français 1 : texte argumentatif	1) J'adresse mon texte à un lecteur – je lui répons en exposant le problème dès le début du texte.	<input checked="" type="checkbox"/>		2/2	
	2) J'exprime clairement mon opinion pour commencer .	<input checked="" type="checkbox"/>		2/2	
	3) Je fournis au minimum 3 arguments .	<input checked="" type="checkbox"/>		6/6	
	4) Dans mon texte apparaissent 2 avantages et 1 inconvénient .	<input checked="" type="checkbox"/>		2/2	
	5) Je rédige une conclusion pour clore mon texte afin de convaincre mon lecteur du choix à faire.	<input checked="" type="checkbox"/>		3/4	
	6) Je m'assure que mon texte soit signé .	<input checked="" type="checkbox"/>		1/1	
	7) Pas de longue phrase .	<input checked="" type="checkbox"/>		0/1	
	8) J'organise mon texte au moyen des organisateur s comme par exemple : tout d'abord, par ailleurs, ensuite, pour terminer.	<input checked="" type="checkbox"/>		4/6	
Français 2 : structuration	9) Je souligne en gris les mots pour lesquels j'ai des doutes. Je vérifie l'orthographe dans un ouvrage de référence.	<input checked="" type="checkbox"/>		1/2	
	10) Je souligne les formes verbales en rouge puis les vérifie au moyen de mon tableau de conjugaison.	<input checked="" type="checkbox"/>		1/2	
	11) Je soigne la propreté, l'écriture et la ponctuation dans mon travail.	<input checked="" type="checkbox"/>		1/2	
Barème					
6	5	4	3	2	1
30-27	26-23	22-19	18-13	12-7	6-1
Commentaires :					

Annexe 2 : Exemple d'entête en mathématiques (élèves de 9-10 ans)

Mes objectifs :		J'auto-évalue mon travail en coloriant la case : (vert : je suis sûr de moi orange : je doute)			
Points Attribués	Objectifs				
/5	• J'effectue des additions et soustractions avec retenue,	😊 😐			
/2	• Je rédige correctement un nombre qui m'est dicté oralement,	😊 😐			
/3	• Je complète une bande numérique,	😊 😐			
/12	• Je résous des problèmes en utilisant correctement les additions et les soustractions,	😊 😐			
/5	• Je me repère dans un plan.	😊 😐			
Barème :					
6	5	4	3	2	1
27-25	24-22	21-18	17-15	14-8	7-1