



REVUE FRANÇAISE
DE
PÉDAGOGIE
recherches en éducation

ÉCOLE NORMALE SUPÉRIEURE DE LYON
Institut français de l'Éducation

Évaluer la formation pratique des enseignants: enjeux, défis et propositions

Référence : André, B. (2013). Évaluer la formation pratique des enseignants : enjeux, défis et propositions. *Revue française de pédagogie* (184), 5-14.

Résumé

L'évaluation des pratiques enseignantes, et plus particulièrement celles des enseignants en formation lors de stages, recèle des difficultés réelles que l'introduction de référentiels de compétences n'a pas atténuées. Pourtant, l'enjeu est particulièrement important, en regard tant du développement professionnel des enseignants que des procédures de *rendre compte* de la part de divers acteurs sociaux. Comment envisager un dispositif permettant de porter un jugement professionnel valide sur la compétence à enseigner ? Cet article tente de faire un inventaire de difficultés auxquelles sont confrontés ceux qui doivent penser de tels dispositifs.

Descripteurs

Formation des enseignants, formation en alternance, formation pratique, évaluation des enseignants, qualification.

Introduction

L'évaluation des pratiques enseignantes a déjà fait l'objet de nombreuses recherches, avec des conclusions allant dans deux directions opposées. Un premier courant, se basant sur les approches *processus-produit* ou la *School Effectiveness Research*, affirme qu'il est possible d'identifier les bonnes pratiques assimilables à un enseignement efficace. L'évaluation consisterait donc à repérer ces pratiques dans l'agir d'un enseignant de manière à qualifier et quantifier sa performance, voire à identifier ses compétences (Bissonnette, Richard, & Gauthier, 2005; Teddie & Reynolds, 2000; Townsend, 2007). D'autres chercheurs mettent en avant la grande variabilité des pratiques chez le même enseignant, et doutent que l'agrégation de multiples bonnes pratiques produise forcément de l'efficacité (Bru, 2002 ; Crahay, 1989).

Face à ces divergences se pose la question de l'identification d'un fondement sur lequel construire la compétence à enseigner et l'évaluer. Après une revue de littérature mettant en évidence la difficulté à formuler des indicateurs de la compétence à enseigner, nous faisons quelques propositions pouvant guider la mise en place d'un dispositif d'évaluation de la formation pratique des enseignants au cours des stages.

Les réflexions qui suivent ont été dans un premier temps suscitées par l'accompagnement de formateurs de terrains dans des dispositifs d'analyse de pratiques, ensuite par l'analyse de bilans de formation de formateurs de terrain à l'issue de leur formation, enfin par la participation à deux reprises au dispositif d'*assessment* de nouveaux enseignants dans une université californienne (*Performance Assessment for California Teachers, PACT*). Elles participent d'une réflexion plus générale sur l'évaluation de la formation pratique à la Haute école pédagogique à Lausanne.

L'identification des pratiques souhaitables

De nombreuses recherches, regroupées sous le type *processus-produit*, ont tenté d'identifier les pratiques enseignantes efficaces, pouvant déboucher sur des standards tant pour la formation que pour l'évaluation. Dans une note de synthèse sur les effets-écoles et les effets-maîtres, Bressoux (1994) relève que « l'effet classe provient pour une grande partie de l'enseignant lui-même » (p. 96). Il souligne six facteurs explicatifs des effets-maîtres : la correspondance entre ce qui est enseigné et ce qui est testé ; le temps alloué à l'apprentissage et celui où les élèves sont réellement engagés dans la tâche proposée ; le taux de réponses exactes aux questions posées par le maître ; ses attentes ; la qualité de ses feedbacks et enfin la structuration des activités. Il relève en outre deux éléments importants : d'une part le type d'enseignement (frontal, de groupe ou individualisé) ne semble pas un facteur clairement discriminant, et d'autre part la combinaison des facteurs est plus importante que chacun pris isolément. De leur côté, dans une publication plus récente, Hattie (2009) et son équipe ont examiné 800 méta-analyses portant sur près de 50'000 recherches en vue d'identifier les variables ayant un impact positif,

neutre ou négatif sur les apprentissages des élèves. Les variables mises en évidence ont été regroupées par ces auteurs dans les catégories : contributions des élèves, de la maison, de l'école, de l'enseignant, des curricula et des approches de l'enseignement, et les effets relevés sur les apprentissages ont été quantifiés à partir des écarts-types constatés relativement à la progression des élèves. La variété des recherches et des méta-analyses ne facilite pas toujours l'interprétation des résultats. Nous citerons cependant les 8 facteurs relevés par Hattie comme produisant le plus d'impact dans la progression des élèves : la qualité de l'enseignement¹ ; l'enseignement réciproque ; la relation enseignant-élève ; la présence des feedbacks ; le soutien à la verbalisation par les élèves des questions et apprentissages ; les stratégies métacognitives ; la pédagogie de type « *Direct Instruction* » et la pédagogie de maîtrise (Hattie, 2009, p. 244).

De leur côté, Gauthier et son équipe (Gauthier, 1997) ont tenté d'élargir l'approche processus-produit, incorporant des dimensions issues des approches cognitivistes et interactionnistes-subjectivistes pour produire un modèle qu'ils caractérisent comme éclectique. Ils tentent ainsi de « saisir ce qui se passe réellement dans la classe et les effets de l'enseignement sur les élèves » (p. 125). Mais même en ajoutant des dimensions issues de différents modèles, force est de constater que le passage du processus agi par le maître au produit montré par l'élève repose sur des raccourcis et des implicites (Bru, Altet, & Blanchard-Laville, 2004) impliquant de sérieuses difficultés lorsqu'il s'agit de se saisir de ces résultats pour fonder une évaluation des pratiques enseignantes.

Pour les détracteurs des approches *processus-produit*, l'enseignement de qualité ne saurait être décrit par une agrégation de bonnes pratiques souvent sans lien formel entre elles. Les raisons invoquées sont multiples. La première est d'abord une affaire de contexte : isoler les bonnes pratiques du réseau complexe des relations qui l'organisent paraît analytiquement et pratiquement insurmontable (Darling-Hammond, Wise, & Klein, 1999 ; Dumay, 2011). C'est ainsi que Clanet (1997) souligne le nombre important de dimensions impliquées dans les interactions liées aux pratiques enseignantes, ces dimensions faisant système. On ne peut prendre un bout de pratique ici, un autre là, et espérer quelque chose qui fonctionne et fasse sens. De plus, le contexte des pratiques est marqué fortement par leur triple dimension historique : l'histoire de la classe, l'histoire de l'enseignant, et l'histoire de la relation entre la classe et l'enseignant. Ces dimensions ont contribué à des ajustements réciproques, des négociations souvent informelles ayant marqué, voire sélectionné et sédimenté des pratiques.

De plus, il s'avère que la variabilité des pratiques chez le même enseignant mais dans différentes classes, est importante, parfois plus que la variabilité entre enseignants (Altet, Bressoux, Bru, & Leconte-Lambert, 1994-1996 ; Crahay, 1989). Ce résultat est particulièrement cohérent avec les travaux montrant l'impact différencié d'une intervention pédagogique selon les caractéristiques des élèves auxquels on s'adresse (Guimond, 2004). D'autres auteurs ont montré que certaines pratiques effectives lorsqu'elles étaient utilisées avec modération, devenaient contre-productives lorsque l'on y recourait intensivement (Peterson & Kauchak, 1982 ; Soar, 1972) ou à mauvais escient (Coker, Medley, & Som, 1980 ; McDonald & Elias, 1976). À cette difficulté s'ajoute le fait que certains effets de l'enseignement ne se manifestent qu'à long terme, rendant particulièrement ardue l'identification des pratiques efficaces (Perrenoud, 2005 ; Prost, 1986).

Deux derniers points méritent d'être soulevés. Tout d'abord, lorsqu'au travers d'une recherche *processus-produit*, l'on cherche à identifier les processus d'enseignement pertinents, il s'agit de s'interroger sur le type de produit valorisé. Carrette (2008) montre que les caractéristiques identifiées chez les enseignants efficaces dépendent des contenus et de la forme des épreuves soumises aux élèves, introduisant un biais potentiel dans les résultats de ces recherches. Enfin, et nous terminerons par là, peut-on asseoir la prescription et l'évaluation d'une pratique pédagogique sur le seul constat d'une cooccurrence entre cette pratique pédagogique et les résultats des élèves ? Ne faut-il pas une intelligibilité minimale des processus impliqués, pour éviter des désillusions lors de la généralisation éventuelle des pratiques identifiées (Bru, 2002) ?

Ces questions mettent en évidence la difficulté à formuler des repères pour asseoir une évaluation des pratiques, lorsqu'il s'agit, par exemple, d'évaluer un étudiant en stage lors de sa formation à l'enseignement. Les pratiques actuelles d'évaluation se centrent sur deux instruments, utilisés parfois conjointement : les référentiels de compétences et les portfolios. Nous les aborderons successivement pour en voir quelques limites et intérêts.

Le recours à un référentiel de compétences

L'accent mis sur la professionnalisation des enseignants s'est accompagné de l'introduction de référentiels de compétences destinés à guider l'élaboration des programmes de formation (Lessard, 2009). De tels référentiels ont pu jouer un rôle, d'importance variable il est vrai, dans le développement d'un « langage commun pour parler de l'activité professionnelle de l'étudiant » (Lessard, 2009, p. 130) entre praticiens, formateurs de terrain, formateur du Centre, et le cas échéant les politiques. Mais ces

¹ Hattie définit la qualité comme étant celle reconnue comme telle par les élèves.

référentiels de compétences peuvent-ils en tant que tels asseoir une évaluation consistante ? Des doutes sont permis (André, 2010 ; Crahay, 2006 ; Paquay, 2006). Une première raison est à chercher dans leur niveau de généralisation. Si l'on prend par exemple la première compétence commune au référentiel vaudois et québécois : « Agir en tant que professionnel critique et porteur de connaissances et de culture », cet énoncé ne dit rien sur ce qui, dans une situation singulière, doit être observé pour pouvoir émettre un jugement positif sur la maîtrise de cette compétence. C'est pour cela que tout référentiel suppose d'être complété par d'autres énoncés permettant de poser, autant que faire se peut, quelques repères pour l'évaluation. Mais on assiste souvent à une multiplication d'énoncés divers et souvent hétérogènes. Si le recours à la notion de compétence présente l'avantage d'éviter une perte de sens par la décomposition de l'agir enseignant en micro-unités, cette atomisation est finalement reportée au niveau des indicateurs, faisant ainsi perdre le bénéfice escompté (F.-M. Gérard & Van Lint-Muguerza, 2000).

Une seconde difficulté s'enracine dans la forte composante relationnelle de l'enseignement (Cifali, 1994 ; Tardif, 1993). Le geste professionnel réussi n'est pas seulement lié à la technicité du métier, mais se tisse dans l'intersubjectivité de la relation, se fondant sur la confiance et la reconnaissance (Sonntag, 2007). Cette dimension, tout en étant essentielle, est difficile à objectiver et à exprimer en utilisant le genre littéraire qui est celui des référentiels. Il s'agit de lire au-delà de l'expertise instrumentale des dimensions qui ne se laissent pas traduire facilement dans des grilles d'observation ou des niveaux de maîtrise. Ainsi les référentiels explicitent des aspects certes importants, tout en laissant implicites d'autres dimensions peut-être plus fondamentales. Que dire, par exemple, de la *présence* en classe ? Cette dimension est constitutive d'un bon enseignement (Richoz, 2009 ; Runtz-Christan, 2000), mais échappe totalement aux énoncés des référentiels.

Dans une autre direction, Berliner (2005), s'appuyant sur les analyses de Fenstermacher et Richardson (2005) distingue trois composantes dans un enseignement de qualité : les actes d'enseignement (définir, démontrer, modeler, expliquer...), les actes psychologiques de l'enseignement (motiver, encourager, évaluer...) et les actes moraux (faire preuve de courage, montrer de la tolérance, de l'empathie, du respect...), sans lesquels un bon enseignement ne peut exister (Hansen, 2001). Si des référentiels se risquent à aborder partiellement cette troisième composante, par exemple sous la forme de l'énoncé : « Agir de façon éthique et responsable dans l'exercice de ses fonctions »², les niveaux de maîtrise associés éludent les enjeux et leur généralité ne se révèlent pas d'une grande aide quant à l'évaluation.

Si l'évaluation ne peut se passer d'une référence à l'aune de laquelle confronter les pratiques observées, les éléments relevés montrent qu'il reste encore du travail à accomplir pour formuler des énoncés plus satisfaisants.

Portfolio et écriture réflexive

De nombreuses formations d'enseignants intègrent actuellement l'utilisation de portfolios ou de dossiers d'apprentissage dans les dispositifs d'évaluation. Issus des métiers artistiques (les « books » des photographes, architectes, etc.) les portfolios ont migré dans les milieux de l'éducation avec une transformation notable. Alors qu'à l'origine ils étaient constitués d'œuvres montrables parce que réalisées sur un support physique, les portfolios doivent maintenant mettre en évidence cette œuvre immatérielle, non montrable qu'est l'enseignement (André, 2012). Ce ne sont donc plus que des supports de l'œuvre (préparation de leçons par exemple) ou des traces secondes (productions d'élèves), voire tierces (analyse et discours sur des traces secondes) qui peuvent figurer dans le portfolio. Dans l'enseignement et la formation, ces dispositifs ont en commun d'inclure une autoévaluation et une réflexivité visant une subjectivation des savoirs (Vanhulle, 2009), et leur mobilisation dans la pratique présente et, on l'espère, ultérieure.

Ces outils peuvent-ils être utilisés comme partie du dispositif de certification ? Cette option soulève plusieurs critiques qui nous paraissent importantes. Premièrement, lorsque le portfolio est un objet sur lequel se fonde une évaluation, il est à craindre que ne se développent des stratégies de façade : « il ne s'agirait alors pas tant d'être réflexif que de montrer des signes — des indicateurs — de sa réflexivité, en réponse aux attentes — supposées ou réelles — des formateurs et du jeu institutionnel » (Huver & Cadet, 2010, p. 128), la compétence révélée étant alors principalement discursive (Presse, 2010). Si l'on ajoute la difficulté à rendre compte de l'ordinaire, de la banalité quotidienne, plus difficile à appréhender que l'extraordinaire, l'évènementiel qui semble dans un premier abord plus « porteur », plus saturé de sens, et donc plus à même de rendre compte de la compétence de celui qui se dit. Cela peut conduire à des fabulations, à des fictions construites à partir de faits scolaires disjoints, vécus ou seulement entendus, dont le rapprochement permet de fabriquer un événement fictif propice à l'exposition de soi et l'analyse... et souvent bien vu des examinateurs (Losego, 1999). Dans une recherche visant à examiner l'éventuelle corrélation entre les compétences observées et les énoncés réflexifs dans des portfolios, Michaud (2012) note que si une telle corrélation peut être mise en évidence sur le corpus qu'ils examinent, l'écriture

² Compétence commune aux référentiel québécois et vaudois.

réflexive ne correspond pas forcément à un agir efficace et à une compétence atteinte.

L'on sait, depuis les travaux de Ricœur (1983-1985), que toute écriture biographique a une composante fictionnelle, due à la mise en intrigue des événements vécus afin de construire du sens. L'illusion biographique, pour reprendre l'expression de Bourdieu (1986), serait donc de prendre pour argent comptant les propos du portfolio, comme exprimant des compétences professionnelles avérées. L'intérêt d'un portfolio en tant qu'outil permettant de construire du sens, de s'appropriier son parcours, d'explorer ses logiques d'action (Heinich, 2010) demeure. Mais son utilisation comme évidence d'une compétence prête à débat.

Des pistes à explorer

Les lignes précédentes montrent la difficulté à évaluer la pratique. Entre la normalisation du « bon enseignant » ou des « *best practices* » et une subjectivité fermée sur elle-même, il s'agit d'explorer la possibilité de construire des consensus dont les limites temporelles et sociétales soient reconnues, étant entendu que l'enseignement, comme d'ailleurs les autres métiers de l'humain, est déterminé par des « pratiques discursives socialement et politiquement marquées, relatives à des finalités, des méthodes et des objets éducatifs » (Mehran, Ronveaux, & Vanhulle, 2007, p. 8).

Construire des standards locaux et provisoires

Sans une forme de standard, il est difficile de dépasser une évaluation intuitive. L'oxymore de l'intertitre ci-dessus vise à mettre en évidence d'une part l'impossibilité de construire des standards universels et la nécessité de disposer d'une référence permettant l'évaluation. La seule voie possible est donc d'admettre le caractère local et provisoire de tels standards. Il s'agit de trouver une bonne granularité des énoncés pour éviter deux pièges symétriques. Le premier est une formulation tellement générale, comme : « *Maîtriser les disciplines et avoir une bonne culture générale* »³, que l'on ne peut que se rabattre sur une évaluation intuitive pragmatique. Le second est d'adopter une formulation si précise qu'elle enferme l'évalué dans un corset prescriptif inadéquat en regard du contexte de sa pratique. Il s'agit en fait de limiter les variations entre les évaluations effectuées par différents acteurs, de manière à préserver leur légitimité en produisant une sorte « d'objectivation relative ».

Cette objectivation relative des résultats de l'évaluation passe notamment par la transformation des standards en échelle d'évaluation. Pour mémoire, il existe deux sortes d'échelles d'évaluation : les échelles uniformes et les échelles descriptives (Scallon, 2004). Les premières sont du type : « Excellent niveau de maîtrise, très bon niveau de maîtrise, bon niveau de maîtrise, niveau de maîtrise satisfaisant, niveau de maîtrise passable, niveau de maîtrise insuffisant ». Leur usage dans la formation des enseignants s'avère délicat, car, en définitive, qu'est-ce qui distingue, par exemple, un excellent niveau de maîtrise d'un très bon niveau de maîtrise ? Celui qui cherche à réfléchir sur ces échelles au-delà de l'idée générale de progression est rapidement dans l'embarras pour fixer un seuil délimitant deux niveaux différents de manière fiable et reproductible. Lorsque l'évaluateur doit assumer ce manque de repère ou la valeur des seuils dans les échelles proposées, cela induit pour lui un coût subjectif important, spécialement lorsqu'il y a échec potentiel. Pour éviter ce coût, certains évaluateurs procèdent à un rabattement de l'évaluation sur l'utilisation de quelques critères très factuels, mais peu intéressants, ou font preuve d'un biais d'indulgence permettant d'éviter les conflits d'interprétations et le risque de devoir se justifier, et renvoyant le cas échéant à d'autres lieux, à d'autres temps, à d'autres évaluateurs, la confrontation de l'étudiant à un éventuel échec. Réciproquement, du côté de l'étudiant évalué, il peut devenir difficile de développer une autoévaluation pertinente si les repères sont inconsistants au fil des différents stages.

Alors que les échelles uniformes sont identiques quel que soit l'item évalué, une échelle descriptive décrit de manière plus ou moins complète et univoque ce qui doit être observé pour évaluer un item déterminé ainsi que les seuils à franchir pour attribuer telle ou telle notation. Par exemple, le *Performance Assessment for California Teachers* développé à l'Université de Stanford (Pechione & Chung, 2006) utilise différentes échelles descriptives dont voici un exemplaire, relatif au standard « L'enseignant utilise les feedbacks pour développer les apprentissages des élèves » et qui interroge la qualité des commentaires adressés aux élèves (The Pact Consortium, 2012) :

³ Référentiel de compétences français pour les professeurs (arrêté du 12 mai 2010)

Niveau 1	Niveau 2	Niveau 3	Niveau 4
<ul style="list-style-type: none"> Les commentaires sont généraux et donnent peu d'indications permettant de progresser relativement aux objectifs d'apprentissage. OU <ul style="list-style-type: none"> Les commentaires contiennent des inexactitudes significatives. 	<ul style="list-style-type: none"> Les commentaires sont opportuns et identifient ce qui a été bien fait ainsi que les domaines d'amélioration liés aux objectifs d'apprentissage. 	<ul style="list-style-type: none"> Des commentaires spécifiques et en temps opportun permettent à l'élève de comprendre ce qu'il a bien fait, et fournissent des indications pour progresser. 	<ul style="list-style-type: none"> Des commentaires spécifiques et en temps opportun permettent à l'élève d'analyser sa propre performance. Les commentaires montrent chez l'enseignant une bonne compréhension de la progression de chaque élève en relation avec les objectifs d'apprentissage et avec la maîtrise du langage.

Elementary Mathematics Rubrics 2012-2013

© 2012 The PACT Consortium (TdA)

L'usage des échelles descriptives comporte aussi son lot de difficultés. En particulier, la difficulté de définir ce qui est véritablement à observer et évaluer, de clarifier les attentes et d'éviter l'inconstance ou l'arbitraire persiste. Mais alors que les échelles uniformes font peser le poids du doute sur l'évaluateur, les échelles descriptives ont le mérite d'affronter le problème dès leur conception. Elles obligent à identifier ou à construire un consensus entre chercheurs, formateurs et praticiens. Ce travail ne peut faire l'impasse sur une méthode itérative entre conception et utilisation, de manière à arriver, après plusieurs étapes, à des énoncés discriminants entre niveaux, compréhensifs pour tous les formateurs, et, autant que possible univoques. Pour assurer cela, il est important que ce travail de formulation réunisse des formateurs du Centre de formation et du terrain, ainsi que chercheurs.

Multiplier les triangulations

La métaphore de la triangulation, qui émane des pratiques des géomètres, décrit en sciences sociales le croisement de données de différente nature en vue d'améliorer l'assise et la validité de résultats de recherche (Mathison, 1988; Webb, Campbell, Schwartz, & Sechrest, 1966). Mottier Lopez et Allal (2010) reprennent ce terme dans leur analyse du jugement professionnel en évaluation, en valorisant le choix et la mise en relation de plusieurs sources d'informations, interprétées au travers de référentiels de différentes natures. En adaptant les propos de ces auteures à l'évaluation des étudiants en stage, on peut considérer les triangulations suivantes :

Triangulations	Démarches possibles
Entre prises de données provenant d'une même méthode	Répétition des mêmes tâches d'observation ou d'enseignement de l'étudiant en stage et mise en évidence de la progression
Entre méthodes différentes	Travail sur des préparations de leçons Travail sur des moments d'enseignement ou des séquences d'enseignement Travail sur des corrections de travaux d'élèves par l'étudiant Utilisation de la vidéoscopie
Entre personnes	Autre formateur de terrain dans le même établissement Visite d'un formateur du Centre, « didactique » ou « transversal » Enregistrement vidéo évalué par un formateur du Centre ou conjointement

Ces différentes possibilités de triangulation, principalement celle du troisième type mentionné, renforcent les contacts entre formateurs de terrain et formateurs du Centre et permettent en plus à chacun de développer son jugement professionnel par la prise en compte du regard de l'autre, de ses problématiques, et des significations potentiellement différentes données aux éléments observés.

Professionnaliser les évaluateurs

Si les deux propositions précédentes portaient sur les instruments et les processus liés à l'évaluation, il nous paraît nécessaire d'aborder la question des évaluateurs et de leur nécessaire compétence à évaluer (Pelletier, 2007). Nos observations montrent qu'un des principaux soucis des formateurs de terrain est la recherche d'objectivité dans leurs évaluations. L'étude de Lehmann (2009) corrobore ce fait. Paradoxalement, ce souci peut s'avérer contre-productif lorsqu'il tente de gommer la subjectivité plutôt que de la prendre en compte et de la travailler, sans pour autant renoncer à la rigueur (Gérard, 2002 ; Hadji, 1997). Vouloir mettre de côté le sujet qui évalue fait courir deux dangers. Le premier est de limiter l'évaluation de l'enseignement à son seul aspect instrumental, plus facilement objectivable et confrontable, au risque de le dénaturer en laissant de côté les autres dimensions relevées précédemment. Le second danger est de laisser libre cours à une forme de subjectivité clandestine (André, 2010 ; Cordonier, 2010), à l'oeuvre sans être reconnue, projetant sur la situation les méandres de sa propre histoire.

La rigueur est à considérer tout au long du processus d'évaluation, que ce soit dans le choix des informations pertinentes à considérer et leur triangulation ; l'utilisation des standards ; l'interprétation et comparaison des données et des standards ainsi que la prise de décision qui s'ensuit. Ce processus de jugement professionnel (Lafortune & Allal, 2007 ; Mottier Lopez & Allal, 2010) des acteurs de l'évaluation, permettant de dépasser un jugement personnel idiosyncrasique. On peut se référer aux travaux de Latour (2001), pour qui la notion d'objectivité est à remplacer par celle de *consensus intersubjectif*, qui implique un accord sur l'objet observé, sur le dispositif de mesure et la confrontation des résultats de mesure dont on veut rendre compte ou aux travaux d'Habermas (1987) concernant l'agir communicationnel. Deux axes de développement et de formation nous paraissent particulièrement nécessaires. Le premier s'articule autour du processus de référentialisation (Figari, 1994) des pratiques enseignantes. Il ne suffit pas d'être capable de donner un sens aux énoncés d'un référentiel ; l'enjeu est aussi de comprendre sa construction, son contexte, le projet qu'il sous-tend et ses limites. Il s'agit autant de construire un rapport critique par rapport à ses propres pratiques vécues comme référentiel implicite, que par rapport aux standards proposés. Le second a trait à la prise en compte de la dimension subjective propre à toute évaluation, de manière à permettre aux évaluateurs de développer leur jugement professionnel avec confiance et rigueur. Ce sont des visées certes ambitieuses, mais à la mesure des défis que pose l'évaluation de la pratique.

Conclusion

L'évaluation des pratiques enseignantes, en particulier celles mises en œuvre lors des stages dans les dispositifs en alternance, révèle des défis importants. Les habitudes et les dispositifs en place ne rendent pas toujours justice à la complexité de cette évaluation, laissant trop souvent place à une évaluation intuitive pragmatique se voulant objective. L'un des enjeux est la qualité de la formation initiale des enseignants. Alors que les enseignants en place font face à la complexification et l'intensification de leur activité (Easthope, C. & Easthope, G., 2000 ; Maroy, 2005), il s'agit de préparer les futurs enseignants le mieux possible aux défis qu'ils rencontreront. Les rétroactions qui leur sont données lors de leurs stages pratiques s'avèrent un facteur crucial pour l'amélioration de leurs pratiques. Cette évaluation formative a tout autant besoin de repères solides et de rigueur que l'évaluation certificative les validant pour l'entrée dans la profession. Dans cette visée, nos propositions ont trois objets. Le premier est relatif à la construction d'un consensus intersubjectif entre formateurs, chercheurs et praticiens formateurs, sous la forme de standards locaux et provisoires, ayant pour but une certaine objectivation de l'évaluation. C'est un consensus limité et conscient de ses limites, qui remplace la poursuite de l'objectivité par celle de la rigueur dans le processus. La deuxième proposition vise à élargir le champ des informations récoltées et penser leur triangulation, de manière à assurer une certaine assise et validité aux résultats de l'évaluation. Enfin, la dernière proposition vise à professionnaliser les évaluateurs au travers du développement de leur jugement professionnel.

L'un des obstacles rencontrés dans nos réflexions est sans conteste la question du coût de ces propositions. En effet, se former, travailler les prises d'informations et leur triangulation, se concerter, s'approprier des outils de travail, réguler, tout cela demande des ressources tant en matière de temps que de disponibilité cognitive. Et le naturel – ou plutôt le naturalisé – revient vite au galop. Pourtant, les bénéfices escomptés sont sans doute à la hauteur des investissements à consentir : la réflexion sur l'évaluation peut enrichir aussi les visées et les dispositifs de formation, et par là soutenir le développement des compétences des futurs enseignants.

Références

- Altet, M., Bressoux, P., Bru, M., & Leconte-Lambert, C. (1994-1996). Étude exploratoire des pratiques d'enseignement en classe de CE2. *Les dossiers de la DEP* (44 et 70).
- André, B. (2010). Évaluation et travail de la subjectivité dans la formation pratique. In L. M. Bélair, C. Lebel, N. Sorin, A. Roy & L. Lafortune (Eds.), *Régulation et évaluation des compétences en enseignement ; vers la*

- professionnalisation* (pp. 145-158). Québec : Presses de l'Université du Québec.
- André, B. (2013). *S'investir dans son travail : les enjeux de l'activité enseignante*. Berne : Peter Lang.
- Berliner, D. C. (2005). The Near Impossibility of Testing for Teacher Quality. *Journal of Teacher Education*, 56(3), 205-213.
- Bissonnette, S., Richard, M., & Gauthier, C. (2005). Interventions pédagogiques efficaces et réussite scolaire des élèves provenant de milieux défavorisés. *Revue française de pédagogie*(150).
- Bourdieu, P. (1986). L'illusion biographique *Actes de la recherche en sciences sociales*, 62-63, 69-72.
- Bressoux, P. (1994). Les recherches sur les effets-écoles et les effets-maîtres. *Revue française de pédagogie*(108), 91-137.
- Brookhart, S. M., & Loadman, W. E. (1992). Teacher Assessment and Validity : What do we want to know? *Journal of personnel Evaluation in Education*(5), 346-357.
- Bru, M. (2002). Pratiques enseignantes : des recherches à conforter et à développer. *Revue française de pédagogie*(138), 63-73.
- Bru, M., Altet, M., & Blanchard-Laville, C. (2004). À la recherche des processus caractéristiques des pratiques enseignantes dans leurs rapports aux apprentissages. *Revue française de pédagogie*(148), 75-87.
- Carette, V. (2008). Les caractéristiques des enseignants efficaces en question. *Revue française de pédagogie*(62), 81-93.
- Cifali, M. (1994). *Le lien éducatif : contre-jour psychanalytique*. Paris : Presses Universitaires de France.
- Clanet, J. (1997). *Contribution à l'intelligibilité du système enseignement-apprentissage. Stabilisations du système et interactions en contexte*. Université de Toulouse II Le-Mirail, Toulouse.
- Coker, H., Medley, D., & Som, R. (1980). How Valid Are Expert Opinions About Effective Teaching? *Phi Delta Kappan*(62), 131-134.
- Connan, P.-Y., & Vincent, J. (2010). Pratiques innovantes et formations universitaires. professionnalisantes : le Portfolio pour développer, accompagner et évaluer les compétences In G. Baillat, D. Niclot & D. Ulma (Eds.), *La formation des enseignants en Europe. Approche comparative* (pp. 155-171). Bruxelles : De Boeck.
- Cordonier, N. (2010). L'épistémologie du « sujet enseignant », tel qu'il apparaît dans les référentiels de compétences et les activités d'expertise. In L. M. Bélair, C. Lebel, N. Sorin, A. Roy & L. Lafortune (Eds.), *Régulation et évaluation des compétences en enseignement ; vers la professionnalisation*. Québec : Presses de l'Université du Québec.
- Crahay, M. (1989). Contraintes de situation et interaction maître-élèves. *Revue Française de Pédagogie*(88), 67-94.
- Crahay, M. (2006). Dangers, incertitudes et incomplétude de la logique de la compétence en éducation. *Revue française de pédagogie*(154), 97-110.
- Darling-Hammond, L., Wise, A. E., & Klein, S. P. (1999). *A License to Teach. Raising Standards for teaching*. San Francisco : Jossey-Bass Publishers.
- Deum, M., & Vanhulle, S. (2008). Portfolio et démarches socio-réflexives en enseignement et formation. *Cahiers de la Section des sciences de l'éducation. Pratiques et théorie*(116).
- Dumay, X. (2011). La recherche sur l'efficacité de l'enseignement ; transformations politiques et modèle d'évaluation *Les dossiers des Sciences de l'Education*(25), 57-68.
- Easthope, C., & Easthope, G. (2000). Intensification, Extension and Complexity of Teachers' Workload. *British Journal of Sociology of Education*, 21(1), 43-58.
- Fenstermacher, G. O., & Richardson, V. (ZOOS) (2005). On making determinations of quality in teaching. *Teachers College Record*, 107(1), 186-215.
- Figari, G. (1994). *Évaluer : quel référentiel ?* Bruxelles : De Boeck.
- Gauthier, C. (Ed.). (1997). *Pour une théorie de la pédagogie*. Bruxelles : De Boeck.
- Gérard, F.-M., & Van Lint-Muguerza, S. (2000). Quelle équilibre entre une appréciation globale de la compétence et le recours aux critères ? In C. Bosman, F.-M. Gerard & X. Roegiers (Eds.), *Quel avenir pour les compétences ?* (pp. 135-140). Paris : DeBoeck Université.
- Gérard, F. M. (2002). L'indispensable subjectivité de l'évaluation. *Antipodes*(156), 26-34.
- Guimond, S. (2004). Evaluer l'impact des pratiques pédagogiques : perspectives de la psychologie sociale expérimentale. *Revue française de pédagogie*(148), 25-36.
- Habermas, J. (1987). *Théorie de l'agir communicationnel*. Paris : Fayard.
- Hadji, C. (1997). *L'évaluation démystifiée*. Paris : ESF.
- Hansen, D. T. (2001). Teaching as a Moral Activity. In V. Richardson (Ed.), *Handbook of Research on Teaching, Fourth edition* (pp. 826-857). Washington : American Educational Research Association.
- Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London : Routledge.
- Heinich, N. (2010). Pour en finir avec "L'illusion biographique". *L'Homme*(195-196), 421-430.
- Huver, E., & Cadet, L. (2010). La formation professionnelle des enseignants : réflexivité et évaluation sont-elles compatibles ? . In L. Paquay, C. Van Nieuwenhoven & P. Wouters (Eds.), *L'évaluation, levier du développement professionnel ?* (pp. 117-129). Bruxelles : De Boeck.
- Lafortune, L., & Allal, L. (Eds.). (2007). *Jugement professionnel en évaluation*. Québec : Presses de l'Université du Québec.
- Latour, B. (2001). *L'espoir de Pandore : pour une version réaliste de l'activité scientifique*. Paris : Ed. La Découverte.
- Lehmann, R. (2009). *Évaluer les gestes professionnels des futurs enseignants : l'impossible mission des formateurs en établissement ? Mémoire de master en sciences de l'éducation*. Genève : Université de Genève.
- Lessard, C. (2009). Le référentiel de compétences, un levier de la professionnalisation de la formation ou un simple effet de langage ? In R. Etienne, M. Altet, C. Lessard, L. Paquay & P. Perrenoud (Eds.), *L'université peut-*

- elle vraiment former les enseignants ? (pp. 127-144). Bruxelles : De Boeck.
- Losego, P. (1999). La construction de la compétence professionnelle et sa mesure : la cas des futurs enseignants formés dans les IUFM. *Revue française de sociologie*, XL(1), 139-169.
- Maroy, C. (2005). « Les évolutions du travail enseignant en Europe. Facteurs de changement, incidences et résistances », *Les Cahiers de Recherche en Education et Formation* 42, p. 1-30.
- Mathison, S. (1988). Why triangulate ? *Educational Researcher*, 17(2), 13-17.
- Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition : the realization of the living*. Dordrecht ; Boston [etc.] : D. Reidel.
- Maturana, H. R., & Varela, F. J. (1987). *The tree of knowledge : the biological roots of human understanding*. Boston ; London : Shambhala.
- McDonald, F. J., & Elias, P. (1976). *Executive Summary Report : Beginning Teacher Evaluation Study, Phase II* Princeton: University of Princeton, Educational Testing Service.
- Mehran, F., Ronveaux, C., & Vanhulle, S. (2007). Du principe d'alternance aux alternances en formation des adultes et des enseignants : un état de la question. In F. Mehran, C. Ronveaux & S. Vanhulle (Eds.), *Alternances en formation* (pp. 67-82). Bruxelles : De Boeck.
- Michaud, C. (2012). Le portfolio, quel lien entre les écrits réflexifs et les compétences ? *Mesure et évaluation en éducation*, 35(2), 9-38.
- Mottier Lopez, L., & Allal, L. (2010). Le jugement professionnel en évaluation : quelles triangulations méthodologiques et théoriques ? In L. Paquay, C. Van Nieuwenhoven & P. Wouters (Eds.), *L'évaluation, levier du développement professionnel ?* (pp. 237-250). Bruxelles : De Boeck.
- Paquay, L. (2006). Evaluer l'activité enseignante. In G. Figari & L. Mottier Lopez (Eds.), *Recherche sur l'évaluation en éducation*. (pp. 51-58). Paris : L'Harmattan.
- Paquay, L. (1994). Vers un référentiel de compétences professionnelles de l'enseignant. *Recherche et formation*(30), 7-38.
- Pecheone, R. L., & Chung, R. R. (2006). Evidence In Teacher Education. The Performance Assessment For California Teachers (PACT). *Journal of Teacher Education*, 57(1), 22-36.
- Pelletier, G. (2007). La professionnalisation du personnel d'encadrement. In L. M. Bélaïr, D. Laveault & C. Lebel (Eds.), *Les compétences professionnelles en enseignement et leur évaluation* (pp. 131-141). Ottawa : Les Presses de l'Université d'Ottawa.
- Perrenoud, P. (2005). Obligation de compétence et analyse du travail : rendre compte dans le métier d'enseignant. In C. Lessard & P. Meirieu (Eds.), *L'obligation de résultats en éducation* (pp. 207-232). Bruxelles : De Boeck.
- Peterson, K., & Kauchak, D. (1982). *Teacher Evaluation: Perspectives, Practices, and Promises*. Salt Lake City : Center for Educational Practice, University of Utah.
- Presse, M.-C. (2010). Les procédures de Reconnaissance et de Validation des Acquis de l'Expérience : quelles tensions ? quelles difficultés ? In L. Paquay, C. Van Nieuwenhoven & P. Wouters (Eds.), *L'évaluation, levier du développement professionnel ?* (pp. 93-104). Bruxelles : De Boeck.
- Prost, A. (1986). *Eloge des pédagogues*. Paris : Seuil.
- Richoz, J.-C. (2009). *Gestion de classes et d'élèves difficiles*. Lausanne : Favre.
- Ricoeur, P. (1983-1985). *Temps et récit*. Paris : Seuil.
- Runtz-Christan, E. (2000). *Enseignant et comédien, un même métier ?* Issy-les-Moulineaux : ESF.
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Bruxelles : De Boeck.
- Shulman, L. S. (2004). *The Wisdom of Practice: Essays on Teaching, Learning, and Learning to Teach*. San Francisco : Jossey-Bass.
- Soar, R. S. (1972). *Follow-Through Classroom Process Measurement and Pupil Growth*. Gainesville: Institute for Development of Human Resources, College of Education, University of Florida.
- Sonntag, M. (2007). Controverses au coeur de la notion de compétence. In L. Talbot & M. Bru (Eds.), *Des compétences pour enseigner* (pp. 105-117). Rennes : Presses universitaires.
- Stierli, E., Tschopp, G., & Clerc, A. (2010). Le dossier de formation est-il un outil d'intégration et de construction de l'identité professionnelle des futurs enseignants ? In L. M. Bélaïr, C. Lebel, N. Sorin, A. Roy & L. Lafortune (Eds.), *Régulation et évaluation des compétences en enseignement ; vers la professionnalisation* (pp. 115-130). Québec : Presses de l'Université du Québec.
- Tardif, M. (1993). Savoirs et expérience chez les enseignants de métier. In H. Hensler (Ed.), *La recherche en formation des maîtres : détour ou passage obligé sur la voie de la professionnalisation ?* (pp. 53-86). Québec : CRP.
- Teddie, C., & Reynolds, D. (Eds.). (2000). *The International Handbook of School Effectiveness Research*. New-York: Routledge.
- The Pact Consortium (2012). *Elementary Mathematics Teaching Event Candidate Handbook 2012-13*. Disponible en ligne à : http://www.pacttpa.org/_main/hub.php?pageName=Teaching_Event_Handbooks#Handbooks.
- Towsend, T. (Ed.). (2007). *International Handbook of School Effectiveness and Improvement. Review, Reflection and Reframing*. New-York: Springer.
- Vanhulle, S. (2009). *Des savoirs en jeu aux savoirs en "je". Cheminements réflexifs et subjectivation des savoirs chez de jeunes enseignants en formation*. Berne : Peter Lang.
- Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive Measures - Nonreactive Research in The Social Sciences*. Chicago : Rand McNally and Co.