

Approche comparative de deux instruments d'évaluation relatifs à l'aptitude à enseigner

Méliné Zinguinian

Haute école pédagogique du canton de Vaud
meline.zinguinian@hepl.ch

Bernard André

Haute école pédagogique du canton de Vaud
bernard.andre@hepl.ch

Résumé

La question de l'évaluation des pratiques enseignantes est une question complexe, tant au regard de leur référentialisation que de la tension entre sur- et sous-prescription présente dans le processus. Notre recherche a examiné les instruments permettant le recueil et la communication des résultats de l'évaluation de deux systèmes relatifs à la formation des enseignants : le *Performance Assessment for California Teachers* (PACT) aux États-Unis et l'évaluation des stages de formation pratique effectués dans le cadre de la formation des enseignants du secondaire inférieur dans une Haute école pédagogique (HEP) en Suisse. Les résultats mettent en évidence des stratégies différentes pour faire face aux difficultés que présente ce type d'évaluation, et interrogent le rôle des différents acteurs de l'évaluation dans l'ensemble des dispositifs examinés.

Mots clés

Évaluation des pratiques enseignantes, référentiel d'évaluation, jugement professionnel, Suisse, Californie

Summary

The issue of the evaluation of teaching practices is a complex one, both in terms of the process of referentialization and the tension between over- and sub-prescription for the evaluator. Our research has examined the instruments for the collection and communication of the evaluation results about two systems concerning the training of teachers: the *Performance Assessment for California Teachers* in the United States and the evaluation of practical training in a University of Teacher Education in Switzerland. The results show different strategies to cope with the difficulties of this type of evaluation, and question the role of the different actors in the three processes examined.

Keywords

Teaching assessment, assessment frame, professional judgment, Switzerland, California

Pour citer cet article : Zinguinian, M & André, B. (2015). Approche comparative de deux instruments d'évaluation relatifs à l'aptitude à enseigner *Evaluer. Journal international de Recherche en Education et Formation*, 1(3), pp. 73-93.

1. Introduction et problématique

L'évaluation des pratiques d'un enseignant pose un certain nombre de questions importantes. La première est sans aucun doute celle de la référentialisation (Figari, 1994, 2014), c'est-à-dire la constitution et la construction du référent à l'aune duquel les pratiques évaluées sont comparées. La seconde est celle posée par la compréhension, l'interprétation et l'utilisation des énoncés du référent par les évaluateurs¹ qui ne vont pas de soi. L'évaluation, entendue comme exercice d'un jugement professionnel, suppose de multiples décisions. Loin d'une simple compilation de données factuelles, elle implique de multiples choix relatifs aux modalités de recueil des données, de leur interprétation et des inférences quant à leurs significations (Mottier Lopez & Allal, 2010). Si l'évaluation sollicite la subjectivité de l'évaluateur (André, 2010 ; Gérard, 2002 ; Weiss, 1986), le travail évaluatif en lui-même sollicite cette subjectivité pour passer « de ce qui est à faire » à « ce qui est fait » (Leplat, 1997). Notre question de recherche est dès lors la suivante : quels sont les impacts des stratégies et documents institutionnels sur le travail évaluatif des formateurs et *in fine* sur l'investissement subjectif qu'il leur demande ?

Il s'avère que l'investissement subjectif des évaluateurs n'est pas proportionnel à leur marge de manœuvre dans les décisions à prendre au niveau de l'évaluation. En effet, l'exploration de plusieurs dimensions pouvant impacter cet investissement, comme l'interprétation des énoncés et le travail d'opérationnalisation des référents ou la prise en compte de l'évaluateur et du contexte de l'évaluation par le système, montre qu'un système fortement - ou au contraire faiblement cadrant - engendre un investissement subjectif similaire.

Comprendre le travail des évaluateurs dans le cadre de l'évaluation des pratiques des futurs enseignants est à la fois un enjeu de connaissance, mais, d'un point de vue pragmatique, questionne aussi la pertinence, la validité et la fiabilité des évaluations posées. Aux États-Unis, ces aspects ont fait l'objet de nombreux débats et travaux (Berliner, 2005; Darling-Hammond, Wise & Klein, 1999; Mitchell, Robinson, Plake, & Knowles, 2001; Wegener Soled, 1995) qui ont abouti à la formulation de standards nationaux de performance pour les nouveaux enseignants par le *Interstate Teacher Assessment and Support Consortium* (InTASC, 2011), dans le but explicite d'améliorer la qualité de l'enseignement (Mitchell, Robinson, Plake & Knowles, 2001). A la Haute école pédagogique (HEP), le texte de présentation du référentiel en usage commence avec cette affirmation : « *L'évolution de l'école et de la pratique de l'enseignement requiert une formation exigeante et permanente. Dans cette perspective, il importe de fixer les finalités de cette formation en termes de compétences professionnelles à développer* » (HEP, 2013, p. 3). La visée principale de ces instruments et processus est donc bien la qualité de la formation. Interroger ces objets s'avère à nos yeux particulièrement pertinent.

Si l'on en vient au point de vue des savoirs, l'analyse de ces objets, leur compréhension et leur rapport au jugement professionnel nous paraissent importants en relation avec le champ de l'activité évaluative en général. Il s'agit d'abord de comprendre le travail des évaluateurs, non seulement du point de vue opératoire, mais aussi en incluant tout ce qui constitue le *travail* au sens de l'ergonomie de langue française, c'est-à-dire tout ce que fait celui qui travaille pour passer de la tâche prescrite à l'activité réalisée (Leplat, 1997). Il s'agit aussi de comprendre les processus proposés dans ces différentes approches : processus planifiés autant que processus réalisés. Poursuivre la compréhension de ce qui se joue dans l'évaluation en partant des instruments d'évaluation institutionnels, tel est donc l'objet de cette recherche.

¹ Pour des raisons de lisibilité, le texte n'a pas été féminisé. Le masculin généraliste inclut systématiquement les femmes.

Cet article vise à traiter ces questions par le biais de l'analyse des documents et des processus qui encadrent l'évaluation des étudiants en formation initiale à l'enseignement et d'entretiens dans deux contextes différents : le *Performance Assessment for California Teachers* (PACT), à savoir la certification permettant aux nouveaux diplômés d'exercer leur métier dans une école de l'Etat de Californie² (Etats-Unis d'Amérique), dont nous avons étudié l'implantation dans une université, et la formation des enseignants dans une Haute école pédagogique romande. L'analyse a porté sur les étudiants se destinant à enseigner au secondaire inférieur en Suisse, ce qui correspond à la Middle School aux Etats-Unis³. Le choix de ces deux contextes est motivé par les différences importantes qu'ils présentent, permettant de mettre en lumière quelques-uns des enjeux de notre question.

Dans le cadre de notre étude, les tâches d'évaluation des pratiques des enseignants en formation sont balisées par des procédures et des documents. Nous ne nous pencherons pas, dans ce travail, sur l'origine de ces prescriptions, sur leurs auteurs ou les processus qui ont conduit à leur énoncé. Non que ce soit inintéressant, mais d'autres l'ont fait avant nous (Périsset & Bagnoux, 2007 ; Lenoir, 2010) et cela sortirait largement du cadre de notre recherche, sans aider à comprendre le travail des évaluateurs, c'est-à-dire ce qu'ils mettent d'eux-mêmes pour passer de ces prescriptions à l'activité concrète qui est la leur. En revanche, la nature et les objets de ces prescriptions constitueront le coeur de notre réflexion, puisqu'ils ont une influence directe sur le travail des évaluateurs.

2. Cadre théorique

Nous examinerons successivement trois aspects relatifs à notre objet : la référentialisation de l'évaluation des pratiques enseignantes, le passage de la tâche à effectuer à l'activité réelle, avec la notion de travail telle que la propose l'ergonomie de langue française, et l'investissement subjectif qu'un tel passage suppose.

2.1 La référentialisation des pratiques enseignantes

Parmi les recherches portant sur l'évaluation des pratiques enseignantes, un premier courant, se basant sur l'approche processus-produit, soutient la possibilité d'identifier les 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 à quantifier sa performance, voire à identifier ses compétences (Bressoux, 1994 ; Hattie, 2009). Cette approche a l'intérêt d'étudier les pratiques enseignantes au regard des résultats des élèves (Attali & Bressoux, 2002).

Néanmoins, pour les détracteurs de cette approche, l'enseignement de qualité ne saurait être décrit par une agrégation de pratiques considérées comme efficaces, car elles sont souvent sans lien formel entre elles (Clanet, 1997; Darling-Hammond et al., 1999; Dumay, 2011). Cette approche ne tient pas non plus compte de l'histoire de la classe, de l'histoire de l'enseignant et de l'histoire de la relation entre la classe et l'enseignant qui ont contribué à des ajustements réciproques ayant sédimenté des pratiques. De plus, elle ne considère pas la variabilité des pratiques chez le même enseignant, dans différentes classes, qui est parfois plus importante que la variabilité entre enseignants (Altet, Bressoux, Bru & Leconte-Lambert, 1994-1996; Crahay, 1989). De même qu'elle n'envisage pas que certaines pratiques efficaces lorsqu'elles sont utilisées avec modération, deviennent contre-productives lorsque l'on y recourt intensivement (Peterson & Kauchak, 1982; Soar, 1972) ou à mauvais escient

² Le PACT a été développé à l'Université de Stanford et est utilisé dans près de trente universités californiennes.

³ Les élèves ont à ce moment entre 11 et 15 ans.

(Coker, Medley & Som, 1980; McDonald & Elias, 1976). À ces difficultés 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). Ces limites étant posées, il n'en reste pas moins que ce courant de recherche a produit des résultats qui méritent d'être considérés. Ils sont particulièrement intéressants s'ils sont pris comme une orientation, une indication, non comme une liste de pratiques à effectuer pour s'assurer un enseignement de qualité (Hattie, 2009). D'autant plus que les chercheurs participant à l'approche processus-produit ont finalement peu observé les pratiques réelles des enseignants, au profit de la mise en relation des prescriptions - de ce qui était attendu des enseignants ou déclaré par eux - avec les résultats des élèves (Tupin, 2003).

Ainsi la référentialisation des pratiques, c'est-à-dire la construction d'un référentiel, à l'aune duquel comparer les informations recueillies lors de l'évaluation d'un enseignant, se révèle une question qui n'est de loin pas triviale. Les éléments relevés mettent en évidence qu'elle demande de la part de l'évaluateur une intelligence du contexte large de l'activité de celui qui est évalué, autant qu'une humilité certaine dans l'utilisation des résultats partiels dégagés par les recherches et l'absence d'un modèle faisant consensus parmi les chercheurs (Brookhart & Loadman, 1992).

Ces différents éléments mettent en évidence la difficulté de la référentialisation (Figari, 1994, 2004) qu'implique l'évaluation des pratiques des étudiants lors des stages. Si le référent est constitué d'éléments génériques et décontextualisés, la marge d'interprétation des évaluateurs sera grande ; si, au contraire, le référent nomme des gestes professionnels précis et spécifiques, le risque existe de passer à côté, soit de pratiques pertinentes mais qui n'entrent pas dans le cadre des énoncés retenus, soit d'une *gestalt*, à savoir une totalité qui a sa valeur, mais dont tous les éléments du référent ne sont pas présents (Royce Sadler, 2009).

Qui dit référent dit aussi référé, et le passage de l'un à l'autre représente plusieurs obstacles importants liés notamment à la compréhension du référentiel par les évaluateurs et aux traductions que requiert le passage d'une observation de situations réelles à un "énoncé conceptuel décontextualisé" (André, dans ce numéro, pp. 4-5).

2.2 Tâche, activité, travail

Depuis les travaux de Ombredane et Favergé (1955), l'analyse du travail se fonde largement sur les recherches en psychologie et en ergonomie faites durant le 19^e siècle. La distinction entre tâche et activité ou, pour reprendre l'expression de Leplat (1997), entre « ce qu'il y a à faire » et « ce que l'on fait » en est un des résultats les plus notables. Cet écart entre ce qui est à faire et ce que l'on fait est toujours présent, alors même que les procédures et les instructions ont souvent la prétention de baliser précisément l'activité dans les tâches dites *d'exécution*. C'est que, entre la situation idéale décrite dans la prescription et les conditions réelles de l'activité, avec leur variabilité et leurs imprévus, il y a le nécessaire engagement cognitif du travailleur, en prise avec une réalité jamais conforme aux modèles censés circonscrire son activité. De plus, les modèles, les logiques et les attentes des diverses instances qui préfigurent son action sont parfois contradictoires, demandant au travailleur des arbitrages difficiles et parfois lourds de dilemmes éthiques. C'est donc entre tâche et activité que prend place le travail, vu comme « la mobilisation coordonnée des hommes et des femmes face à ce qui n'est pas donné par l'organisation prescrite du travail » (Davezies, 1993). C'est dans cette mobilisation que certains voient l'investissement subjectif (André, 2013 ; Clot, 1999; Dejours, 1993; Jobert, 1999), en le décrivant comme consubstantiel au travail. L'investissement subjectif est nécessaire, parce que sans lui le travail lui-même ne saurait se produire.

2.3 L'investissement et le coût subjectif dans l'évaluation

L'investissement subjectif, au sens de ce que le travailleur met de lui-même pour faire advenir les résultats souhaités, est particulièrement important dans deux cas de figure : lorsque la prescription, c'est-à-dire l'ensemble des éléments qui définissent la tâche, est trop lâche et laisse une très grande marge de manœuvre au travailleur (on parle alors de *tâche sous-prescrite, de déficit de prescription ou de dérégulation*), ou, au contraire, lorsque la prescription est à ce point précise qu'elle se heurte aux variabilités des contextes et des conditions de travail, niant la possibilité autant que la nécessité d'adapter son action pour atteindre les objectifs fixés (*tâches sur-prescrites*) (Falzon et Sauvagnac, 2004).

Sous-prescription comme sur-prescription ont un coût, lié à la renormalisation de l'activité. Dans le cas de l'évaluation, la sous-prescription apparaît lorsque le processus ou le référent est flou, laissant une large place aux interprétations de l'évaluateur, et le laissant porter seul les décisions liées à l'évaluation. La sur-prescription intervient lorsque processus et référent sont à ce point protocolés qu'ils ne laissent plus de place au jugement professionnel de l'évaluateur pour, par exemple, intégrer d'autres informations que celles définies par les prescripteurs. C'est un coût subjectif, par opposé aux coûts objectifs de l'évaluation, que sont le temps ou les ressources matérielles nécessaires pour la mise en place de l'évaluation.

3. Contexte et méthodologie

Afin de saisir la complexité du phénomène qui nous intéresse, l'analyse a porté sur deux contextes différents : une Haute école pédagogique romande et une université californienne. Cette Haute école pédagogique se caractérise par une organisation de l'évaluation de la pratique spécifique, dans le sens qu'elle repose essentiellement sur les épaules du formateur de terrain qui remplit ainsi la double fonction de formateur et d'évaluateur du stagiaire. L'université californienne, pour sa part, se caractérise par sa taille. En effet, il s'agit d'une petite université, ce qui implique que, bien que le système du PACT soit imaginé pour que l'évaluateur ne soit pas le formateur du stagiaire, dans les faits ce n'est pas toujours possible d'aller dans ce sens, principalement dans ce type d'universités qui ne comptent pas forcément plusieurs formateurs par discipline. Ces deux institutions, sans être représentatives de tout ce qui se fait dans les régions où elles sont implantées, illustrent, de par leurs spécificités, une certaine réalité des systèmes étudiés.

Le recueil de données s'est déroulé de 2012 à 2013. Il a tout d'abord consisté en la compilation et l'analyse des documents officiels : les référentiels eux-mêmes, référentiel de compétences professionnelles pour la HEP, *California Teaching Performance Expectations* (TPE) (Commission on Teacher Credentialing, 2013) pour le PACT et les outils de l'évaluation, soit des extraits du document de suivi de stage pour la HEP et des échelles descriptives pour le PACT.

Les standards formulés dans les TPE sont le support de l'évaluation des enseignants en formation pour l'enseignement à la *primary* et à la *secondary school* en Californie⁴. Ils sont communs à toutes les disciplines d'enseignement et servent de base à l'élaboration des échelles descriptives (*rubrics*), élaborées par le Consortium du PACT (2010), spécifiques à chaque discipline et qui elles servent d'outil pour l'évaluation. Dans cette étude, l'analyse porte sur les échelles descriptives pour l'évaluation de l'enseignement en mathématiques.

⁴ Ce qui correspond à l'enseignement à des élèves de 4 à 19 ans environ.

Le référentiel de compétences élaboré par des formateurs de la HEP (2013b) est également commun à toutes les disciplines et aux degrés de l'enseignement primaire et secondaire⁵. Les documents de suivi de stage, élaborés à partir du référentiel de compétences, servent dans ce contexte d'outil pour l'évaluation et sont, contrairement aux échelles descriptives, similaires pour toutes les disciplines. En revanche, ils sont adaptés au stage évalué afin d'inscrire une logique de progression dans la formation et au degré d'enseignement dans lequel le stagiaire est engagé (primaire, secondaire inférieur, secondaire supérieur). Dans cette étude, l'analyse porte sur le document d'évaluation du quatrième et dernier stage de la formation pour l'enseignement au secondaire inférieur (HEP, 2013c).

Parallèlement, nous avons recueilli des informations sur la pratique des formateurs de terrain par l'analyse de travaux de bilan rédigés par ces acteurs à l'issue de la formation qu'ils suivent à la HEP. Pour le PACT, quatorze entretiens semi-dirigés, d'une durée de 45 minutes environ, ont été réalisés dans l'université californienne visitée, auprès de sept étudiants qui venaient de passer le PACT, de trois formateurs de terrain, d'un formateur universitaire engagé dans l'évaluation du PACT, d'un superviseur universitaire et de deux enseignants novices ayant passé le PACT précédemment. Les données ainsi récoltées documentent le processus d'évaluation dans son ensemble.

Nos analyses reposent donc sur une variété de sources que nous avons croisées pour comprendre le travail des évaluateurs en relation avec la nature et les contenus des documents et procédures utilisés.

⁵ Ce qui correspond à l'enseignement à des élèves de 4 à 18 ans environ.

Tableau 1 : Sources des analyses

Établissement	Sources utilisées	Contenu
HEP	Référentiel de compétences professionnelles	Description des 11 compétences professionnelles et énoncé des niveaux de maîtrise attendus en fin de formation.
	Documents de suivi de stage	Intitulés des 11 compétences professionnelles, complétés par les niveaux de maîtrise attendus en fin de formation et des critères d'évaluation pour certaines compétences. ⁶
	Bilans de formation	Bilan de formation et d'activité des formateurs de terrain écrits à l'issue de leur formation.
PACT	TPE	Description des 13 standards présentant les aptitudes attendues d'un enseignant californien.
	Mathematics Rubrics	Description des performances à observer dans la pratique et dans le dossier d'évaluation du candidat, en relation avec les TPE.
	Interviews	Entretiens semi-dirigés portant sur l'expérience vécue du PACT par les acteurs interrogés, sur le processus et les rôles occupés par les acteurs, la perception qu'ils ont des rôles des autres intervenants et celle des avantages et inconvénients de ce système d'évaluation.

4. Résultats

Nous présenterons successivement les éléments relatifs aux acteurs, aux processus et aux référents.

4.1 Les acteurs

Les deux contextes étudiés se présentent de la manière suivante :

Tableau 2 : Présentation des deux systèmes analysés

	HEP	PACT
Evaluation	Evaluation de l'aptitude à enseigner sur la base d'un référentiel de 11 compétences accompagnées de niveaux de maîtrise et d'une échelle uniforme (<i>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</i>).	Evaluation de l'aptitude à enseigner à l'issue de la formation, en utilisant des standards de performance traduits en échelles descriptives (voir tableau 6).

⁶ Les documents identifient *critères* et *niveaux de maîtrise* sans jamais définir explicitement le sens donné à ces termes. Pour notre part, nous utiliserons le terme générique de *descripteur* pour désigner un énoncé décrivant, sous une forme ou sous une autre, ce qui devrait être constaté dans l'activité de l'évalué pour inférer sa compétence.

	HEP	PACT
Acteurs	L'évaluation est sous la responsabilité du formateur de terrain si le stage se déroule de manière satisfaisante. Dans ce cas, il lui revient d'attribuer la note finale du stage. Une co-évaluation impliquant des formateurs de la HEP ainsi que de l'unité responsable de la formation pratique a lieu en cas de difficulté.	L'évaluation se fait par un membre de la communauté académique, y compris l'attribution de la note finale, sur la base d'un dossier comprenant un extrait vidéo et des textes documentant la séquence filmée. Les évaluateurs sont préalablement formés et utilisent des vidéos de référence pour s'accorder sur l'évaluation (processus de <i>benchmarking</i>).

Les deux systèmes ont un objectif commun plus ou moins explicite : évaluer, de manière rigoureuse, l'aptitude à enseigner des futurs enseignants pour favoriser un enseignement de qualité. Néanmoins, les chemins empruntés sont divers et les responsabilités des acteurs impliqués également :

Tableau 3 : Responsabilités des acteurs de l'évaluation

	HEP	PACT
Évalué	Mettre en œuvre un enseignement en accord avec les compétences visées, les attentes du formateur de terrain et les niveaux de maîtrise attendus à la fin de chaque stage. Il n'est pas attendu de lui une participation directe à son évaluation (portfolio, auto-évaluation écrite, etc.).	Mettre en œuvre un enseignement en accord avec les performances attendues. Enregistrer une ou plusieurs séquences vidéo montrant toutes les facettes des performances attendues. Choisir la séquence correspondant le mieux aux attentes du système. Rendre compte par écrit de son enseignement.
Évaluateurs	Interpréter la description proposée des compétences, les niveaux de maîtrise et les critères. Recueillir les éléments jugés significatifs en fonction de ces interprétations. Comparer les intitulés des compétences, critères et niveaux de maîtrise attendus avec les éléments recueillis, qualifier les compétences de l'étudiant en utilisant une échelle uniforme. Rendre compte de son évaluation au stagiaire.	Vérifier la conformité des données recueillies par l'étudiant. Accorder ses interprétations des énoncés et des observations à celles des chercheurs et des pairs par des procédures de <i>benchmarking</i> . Comparer les descripteurs et les éléments recueillis, les qualifier en utilisant les échelles descriptives. Introduire le résultat dans le système de gestion informatique.

Système	<p>Élaborer le référentiel de compétences et les niveaux de maîtrise.</p> <p>Intituler et décrire les compétences, définir les niveaux de maîtrise et les critères.</p> <p>Former les formateurs de terrain pour exercer leur fonction, y compris celle d'évaluation de la pratique.</p> <p>Mettre en place une procédure de co-évaluation en cas de difficulté en stage.</p>	<p>Élaborer les standards de performance et les échelles descriptives.</p> <p>Mettre en place des procédures comparatives permettant aux évaluateurs d'accorder leurs interprétations et d'étalonner leurs qualifications des observations.</p> <p>Contrôler les évaluations faites par échantillonnage pour vérifier leur conformité au modèle et réguler les procédures.</p>
----------------	---	--

Cette comparaison met en évidence différentes stratégies utilisées dans la répartition des tâches liées à l'évaluation et dans la régulation des dispositifs. Ainsi, dans le PACT, l'évalué joue un rôle central dans la récolte d'informations au travers de la sélection d'une séquence vidéo qui sera un élément clé dans son évaluation (Pechone & Chung, 2007). Les interviews de plusieurs candidats montrent que cette tâche est particulièrement ardue, tant du fait des contraintes techniques (assurer la visibilité de son action dans l'enregistrement) que de la nécessité de montrer plusieurs facettes importantes de son aptitude à enseigner dans un temps restreint et des circonstances imprévisibles.

"[...] there's a lot of challenges that you don't think about with the video clip. You know I had my lesson all set up, I had it all planned, and then there's a portion of clip I wanted to show where somebody's talking next to the camera and you can't hear it that well so I had to edit that out or choose a different clip! So choosing the video clips was a little challenging just because of all the little things or you know, am I standing off the frame, or setting it all up [...]." (Extrait d'entretien avec un enseignant novice ayant passé le PACT précédemment)

"So for me I recorded everything just to be sure. And then it was difficult picking which one I want to submit to be the best – which one fulfilled everything that needed to be fulfilled - and answered all the questions." (Extrait d'entretien avec un étudiant venant de passer le PACT)

A contrario, l'évaluation à la HEP (2013c) peut s'étendre dans le temps et le relevé des actions de l'étudiant est à la charge du formateur de terrain.

Autre différence notable : l'importance, dans le PACT, donnée à une homogénéisation des évaluations au travers de procédures de *benchmarking* dans la formation des évaluateurs et de la régulation du système, procédures absentes à la HEP.

4.2 Les processus

Si l'on compare maintenant les processus, on observe des différences marquantes. Tout d'abord la HEP déclare évaluer des compétences et le PACT des performances. La compétence au sens de la *capacité de mobiliser des savoirs et des savoir-faire constitués, dans l'urgence et l'incertitude*, en situation réelle (Perrenoud, 2001), est par essence inobservable : elle s'infère à partir d'un certain nombre de *performances*, c'est-à-dire d'actions réussies ou non (Roegiers, 2000 ; Scallon, 2004). L'inférence de la compétence permet d'affirmer raisonnablement que le candidat est capable de faire face à telle famille de situations.

L'observation de la performance, dans le cadre du PACT, a une moindre ambition : permettre de constater que le candidat a réussi son intervention dans le contexte observé. La différence entre les deux contextes réside donc dans le nombre de performances : une seule pour le PACT, plusieurs, échelonnées sur un semestre, pour la HEP.

Une deuxième différence touche à la restitution des résultats de l'évaluation. A la HEP (2013c), il revient au formateur de terrain qui a accompagné le stagiaire de lui faire part de l'évaluation de ses compétences et, au stagiaire, de remettre à l'institution le formulaire d'évaluation signé par les deux parties. Concrètement, au terme du stage, et bien qu'aucune contrainte à ce sujet ne figure dans les documents officiels analysés, le formateur rencontre le stagiaire afin de lui présenter l'évaluation qui a été réalisée et lui faire des commentaires pour la suite de sa formation et de sa carrière. Le poids de la relation directe et de la dimension intersubjective ne peut donc être négligé dans le processus. Dans le PACT en revanche, les résultats sont introduits par l'évaluateur dans le système informatique permettant la gestion de tout le processus, et l'étudiant en prend connaissance en consultant un site web, évitant ainsi tout contact direct entre l'évaluateur, qui reste anonyme, et l'évalué.

Enfin, nous terminerons en soulignant la manière différente d'opérationnaliser les référentiels au moment de l'élaboration des outils d'évaluation dans ces deux systèmes, en commençant par présenter ces documents.

4.3 Les référents

4.3.1 Les référentiels

Les référents de l'évaluation sont donc, d'une part, le référentiel de compétences pour la HEP et, d'autre part, les TPE pour le PACT. Ces deux textes dont l'intention est clairement normative, formulent les attentes vis-à-vis d'un enseignant reconnu comme compétent et ont une portée heuristique similaire : « *quand on déclare qu'un enseignant est compétent ou incompétent, qu'un professeur d'université est compétent dans son domaine, on utilise le terme de compétence comme un attribut de la personne à partir d'un jugement fondé sur des standards de performance plus ou moins explicites* » (Louis, Jutras & Hensler, 1996, p. 418).

Le tableau 4 met en concordance les énoncés des compétences du référentiel HEP avec ceux des TPE. Il permet ainsi d'identifier les similitudes et les différences tant au niveau du contenu que de la formulation des énoncés.

Tableau 4 : Comparaison entre le référentiel de compétences de la HEP et les TPE

Référentiel HEP	TPE ⁷
1. Agir en tant que professionnel critique et porteur de connaissances et de culture	<p><i>A. Making Subject Matter Comprehensible for Students</i> TPE 1: Specific pedagogical skills for subject matter instruction (teaching English, social studies, mathematics, science, world languages, physical education, art, music)</p>
2. S'engager dans une démarche individuelle et collective de développement professionnel	<p><i>F. Developing as a Professional Educator</i> TPE 12: Professional, legal, and ethical obligations (teaching tolerance, promoting democratic principles, resisting racism and sexual harassment) TPE 13: Professional growth (evaluating teaching practices, using reflection and feedback to improve teaching, increasing subject matter knowledge)</p>
3. Agir de façon éthique et responsable dans l'exercice de ses fonctions	
4. Concevoir et animer des situations d'enseignement et d'apprentissage en fonction des étudiants et du plan d'études	<p><i>D. Planning Instruction and Designing Learning Experiences for Students</i> TPE 8: Learning about students (understanding patterns of adolescent development, encouraging parents to be involved in student learning, understanding how factors such as gender and health can influence student learning) TPE 9: Instructional planning (establishing short- and long-term learning goals, planning and sequencing daily instruction, accommodating student needs through a wide variety of instructional strategies)</p>
	<p><i>C. Engaging and Supporting Students in Learning</i> TPE 4: Making content accessible (incorporating specific strategies that motivate students to learn, meeting student academic learning needs in a variety of ways) TPE 5: Student engagement (ensuring participation of all students, examining multiple points of view, encouraging student cooperation and sharing, using student experience to make instruction relevant) TPE 6: Developmentally appropriate teaching practices (explaining and using practices that are of greatest importance and effectiveness in teaching adolescents such as problem solving, concrete reasoning and abstract thinking, connecting the curriculum to life beyond the classroom)</p>

⁷ Nous avons modifié l'ordre de présentation pour mettre les standards en correspondance avec les compétences relatives aux mêmes champs.

Référentiel HEP	TPE ⁸
5. Évaluer la progression des apprentissages et le degré d'acquisition des connaissances et des compétences des élèves/étudiants	<p><i>B. Assessing Student Learning:</i></p> <p>TPE 2: Monitoring student learning during instruction (determining progress toward learning, pacing of instruction, using questioning strategies, examining student work)</p> <p>TPE 3: Interpretation and use of assessments (using formal and informal assessments, developing multiple assessment measures, using standardized tests, providing feedback on student learning)</p>
6. Planifier, organiser et assurer un mode de fonctionnement de la classe favorisant l'apprentissage et la socialisation des élèves/étudiants	<p><i>E. Creating and Maintaining Effective Environments for Student Learning</i></p> <p>TPE 10: Instructional time (establishing procedures for routine instructional tasks, managing transitions to maximize instructional time)</p> <p>TPE 11: Social environment (promoting a positive, safe environment, creating and using a student discipline plan)</p>
7. Adapter ses interventions aux besoins et aux caractéristiques des élèves/étudiants présentant des difficultés d'apprentissage, d'adaptation ou un handicap	TPE 7: Teaching English learners (applying instructional theories, principles and practices for instruction of English language learners)
8. Intégrer les technologies de l'information et de la communication aux fins de préparation et de pilotage d'activités d'enseignement et d'apprentissage, de gestion de l'enseignement et de développement professionnel	
9. Travailler à la réalisation des objectifs éducatifs de l'école avec tous les partenaires concernés	
10. Coopérer avec les membres de l'équipe pédagogique à la réalisation de tâches favorisant le développement et l'évaluation des compétences visées	
11. Communiquer de manière claire et appropriée dans les divers contextes liés à la profession enseignante	

⁸ Nous avons modifié l'ordre de présentation pour mettre les standards en correspondance avec les compétences relatives aux mêmes champs.

Si le nombre d'énoncés est sensiblement identique dans les deux référentiels (respectivement 11 et 13), il n'y a pas de correspondance systématique entre les deux listes. Nous relèverons en particulier que :

- 4 compétences de la HEP n'ont pas d'équivalent dans les standards (compétences 8 à 11)
- 5 standards (TPE 4 à 6, ainsi que 8 et 9) sont directement en lien avec la compétence 4 renvoyant à l'enseignement lui-même.

Le degré de généralité des énoncés est relativement élevé, bien que les énoncés soient plus détaillés dans la seconde colonne (TPE). Ils sont généraux dans le sens où ils ne mentionnent pas de geste professionnel spécifique, de situations dans lesquelles ils doivent être observés ni la manière de répondre aux attentes décrites (comment planifier un mode de fonctionnement de la classe propice aux apprentissages et à la socialisation des élèves ou favoriser des interactions sociales positives - qu'est-ce qu'une interaction sociale positive, ceci n'est également pas précisé dans les référentiels).

4.3.2 Les outils de l'évaluation

Les outils de l'évaluation diffèrent notablement tant sur leur forme que sur le fond. Si l'on prend l'exemple de l'évaluation d'une compétence ou d'un standard, pour la HEP, il s'agit d'une seule liste d'items à considérer, avec la difficile et délicate question de l'agrégation des résultats de l'évaluation d'items reliés à des dimensions différentes de l'enseignement. Pour le PACT, ce sont deux échelles descriptives qui sont formulées :

Tableau 5 : Compétence 4, niveaux de maîtrise et critère (HEP, 2013c)

Compétence 4 : Concevoir et animer des situations d'enseignement et d'apprentissage en fonction des élèves et du plan d'études
<p>Niveaux de maîtrise (* le cas échéant, critère-s)</p> <ul style="list-style-type: none"> - Analyser et évaluer la matière contenue dans les plans d'études afin de réguler les activités d'enseignement-apprentissage. - Concevoir des activités d'enseignement-apprentissage variées, cohérentes et fondées aux plans didactique et pédagogique, et d'un niveau de complexité permettant la progression des élèves dans le développement de leurs compétences. - Intégrer les activités d'enseignement-apprentissage dans une planification globale. - Guider, par des interventions appropriées, les élèves dans leurs activités d'apprentissage. - Adapter le déroulement de l'activité aux réalités du moment. - Prendre en compte la diversité des démarches des élèves. - Détecter les problèmes d'enseignement-apprentissage qui surviennent et utiliser les ressources appropriées pour y remédier. <p><i>* Organiser et conduire des périodes d'enseignement (45'-90') adaptées au niveau d'apprentissage des élèves.</i></p>

Tableau 6 : Mathematics Rubrics, Last updated 2012
 © 2010 The PACT Consortium.(TdA)

Enseignement : Engager les élèves dans les apprentissages (Mathématiques)			
Comment l'étudiant engage-t-il activement les élèves dans la construction de leur compréhension des concepts, procédures et raisonnements mathématiques? (TPE 1, 5 et 11)			
Niveau 1	Niveau 2	Niveau 3	Niveau 4
<p>Dans le clip, les élèves ont peu d'occasions de s'engager dans des activités leur permettant de développer leur compréhension des concepts, procédures et raisonnements mathématiques.</p> <p>OU</p> <p>Le clip ne met pas en évidence un travail sur la compréhension des concepts, procédures et raisonnements mathématiques.</p> <p>OU</p> <p>La gestion de la classe est problématique et le comportement des élèves est un obstacle pour l'apprentissage.</p>	<p>Le clip met en évidence des stratégies permettant aux élèves de s'engager pour développer leur compréhension des concepts, procédures et raisonnements mathématiques.</p>	<p>Le clip met en évidence des stratégies permettant aux élèves de s'engager activement pour développer leur compréhension des concepts, procédures et raisonnements mathématiques.</p> <p>Ces stratégies tiennent compte des caractéristiques des élèves, de leurs besoins particuliers concernant l'apprentissage et de leur habileté langagière.</p>	<p>Le clip met en évidence des stratégies permettant aux élèves de s'engager activement pour développer leur compréhension des concepts, procédures et raisonnements mathématiques.</p> <p>Ces stratégies sont explicites et mettent en évidence la prise en compte des caractéristiques des élèves, de leurs besoins particuliers concernant l'apprentissage et de leur habileté langagière.</p>

Enseignement Guider les apprentissages (Mathématiques)			
Comment l'étudiant guide-t-il les apprentissages des élèves et interagit-il avec les questions, commentaires et besoins des élèves ? (TPE 2 et 5)			
Niveau 1	Niveau 2	Niveau 3	Niveau 4
<p>Le candidat conduit son enseignement principalement en posant des questions superficielles et évalue les réponses comme correctes ou incorrectes.</p> <p>Les réponses du candidat n'encouragent pas les élèves à réfléchir.</p> <p>OU</p> <p>Les documents ou les réponses du candidat contiennent des erreurs significatives entraînant des incompréhensions chez les élèves.</p>	<p>Le candidat suscite des réponses des élèves qui demandent un raisonnement mathématique ou des stratégies de résolution de problèmes.</p> <p>Les réponses du candidat mettent en évidence sa volonté d'améliorer la compréhension des élèves concernant les concepts, les procédures et les raisonnements mathématiques.</p>	<p>Le candidat suscite des réponses des élèves qui demandent un raisonnement mathématique ou des stratégies de résolution de problèmes.</p> <p>Les réponses du candidat utilisent les inputs des élèves pour guider leur compréhension des concepts, des procédures et des raisonnements mathématiques.</p>	<p><i>Contenu du niveau 3, avec en plus :</i></p> <p>Le candidat suscite l'explicitation des raisonnements mathématiques ou des stratégies de résolution de problèmes des élèves, et utilise ces explicitations pour approfondir la compréhension de tous les élèves.</p>

En mettant en regard quelques caractéristiques de ces deux types d'énoncés utilisés pour l'évaluation, nous obtenons le tableau suivant :

Tableau 7 : comparaison des énoncés liés à l'évaluation dans les deux systèmes

HEP	PACT
Niveaux de maîtrise énoncés sous forme d'activités des enseignants.	Descripteurs énoncés en termes de stratégies ou de comportements attendus de l'enseignant.
Aucun lien direct entre les descripteurs et les niveaux de l'échelle uniforme, aucune définition des seuils entre les niveaux.	Mise en relation de ce qui est observé avec l'échelle de l'évaluation au travers d'une grille (échelle descriptive).
Les énoncés décrivent des intentions, et l'évaluation de leur réalisation est une variable continue.	Les énoncés se veulent <i>evidence-based</i> : le descripteur est observé ou pas, dans une perspective binaire.

Si l'on considère maintenant le travail d'opérationnalisation dont ces outils font l'objet, la HEP stipule que les critères mentionnés pour préciser les compétences ne forment pas, lorsqu'ils existent, une liste exhaustive d'indicateurs à évaluer (HEP, 2013c), ce ne sont d'ailleurs pas véritablement des indicateurs tels que décrits dans la littérature sur l'évaluation

des compétences (Scallon, 2004). Ce faisant, cette institution renvoie aux évaluateurs la responsabilité de tout ou partie du choix des indicateurs qu'ils considèrent comme pertinents, pour pouvoir servir de justificatif lorsqu'ils devront rendre compte de leur évaluation.

Dans ce contexte, les niveaux de maîtrise et les critères servent *in fine* à guider l'évaluation de chaque compétence prise dans sa globalité. La HEP valorise ainsi un jugement global de la prestation du stagiaire. Ceci est renforcé par le fait que les critères utilisés (2013c) restent largement englobants en termes d'actes d'enseignement réalisés par le stagiaire : "Identifier différentes tâches dévolues à l'enseignant" (critère de la compétence 2, HEP, 2013c). Il est à noter que si la HEP encourage un regard global sur les compétences du stagiaire, elle ne renvoie pas intégralement à l'évaluateur le travail d'interprétation du référent, en proposant des niveaux de maîtrise et le cas échéant des critères. Une opérationnalisation minimale est ainsi prise en charge au niveau du système, minimale au regard de ce qui est élaboré dans le PACT.

Le PACT propose deux ou trois éléments à identifier pour chaque niveau de maîtrise (PACT Consortium, 2010). Il ne s'agit donc pas ici de porter un jugement global sur la performance du candidat, mais bien d'identifier, à l'aide de ces descripteurs spécifiques, certains aspects d'un acte d'enseignement : "Le clip met en évidence des stratégies permettant aux élèves de s'engager activement pour développer leur compréhension des concepts, procédures et raisonnements mathématiques" (niveau 3, tableau 6). Les stratégies ne sont certes pas spécifiées, mais les différents niveaux décrits en dessinent les caractéristiques, ce que le critère de la compétence 2 du référentiel HEP, présenté dans le tableau 5, ne propose pas. Au même titre que la HEP n'a pu faire l'économie d'un minimum d'opérationnalisation au niveau du système, le PACT, tout en orientant l'évaluation du candidat vers des éléments précis de son enseignement, ne va pas jusqu'à dicter les gestes concrets à exécuter. Les échelles descriptives ne mentionnent pas, par exemple, les stratégies à employer pour permettre aux élèves de s'engager dans les apprentissages.

5. Discussion des résultats

5.1 Les référentiels à l'épreuve du travail réel des évaluateurs

Cette étude, en soulevant la difficile question de la référentialisation de l'évaluation de l'aptitude à enseigner, a mis au jour le travail de renormalisation effectué à partir des référentiels afin de les opérationnaliser. Passer d'une liste de compétences ou de standards à un instrument permettant l'évaluation a demandé une traduction, un retravail des énoncés. La HEP est celle qui relie le plus étroitement ses grilles d'évaluation au référentiel de compétences professionnelles. Ceci peut être dû au fait qu'elle l'a elle-même choisi et adapté à ses besoins, évitant une réécriture d'un texte formulé par d'autres. Tout au plus, des critères ont été formulés, ainsi que les niveaux de maîtrise attendus en fin de formation. On peut voir, dans le fait que les évaluateurs ont la possibilité d'utiliser d'autres indicateurs, une reconnaissance assumée de l'incomplétude des indicateurs possibles de la compétence, et, en parallèle, une reconnaissance du jugement professionnel des évaluateurs.

Concernant le PACT, il est possible de voir, dans le travail de renormalisation effectué par ses acteurs pour passer des TPE aux échelles descriptives, une visée, *a priori*, de création d'un document exploitable très concrètement par les évaluateurs, et dénué de toute ambiguïté, de manière à assurer la fiabilité et la validité de l'évaluation. Mais ne peut-on lire aussi dans ce remaniement un embarras face à des énoncés forcément simplificateurs en regard de la complexité des pratiques enseignantes ? Tout se passe comme si tout référentiel, quel qu'il soit, et parce qu'il est le fruit de décisions, de valeurs, d'une histoire, d'un contexte, devait forcément être réécrit par ses utilisateurs, pour que ces derniers puissent se l'approprier. À

moins que le référentiel ne soit imposé sans négociation possible, c'est-à-dire sans participation possible pour les utilisateurs au travail de réécriture, ce qui est le cas dans les deux systèmes étudiés. Néanmoins, les utilisateurs, par l'interprétation qu'ils font du contenu de ces documents, participent, même si cela ne peut impacter les référentiels eux-mêmes, à leur renormalisation. En poussant à l'extrême, un référentiel n'est référence que lorsqu'il est formulé par celui qui l'utilise. Ceci nous renvoie directement à la question de l'investissement subjectif dans l'évaluation.

5.2 Vers une reconnaissance du jugement professionnel

Le fait que la HEP ne formule que quelques critères et des niveaux de maîtrise est, comme nous venons de le voir, une reconnaissance du jugement professionnel des évaluateurs, mais en même temps une délégation d'un surcroît de responsabilités. Cette large place accordée au jugement professionnel de l'évaluateur incite à l'investissement subjectif de ce dernier dans son activité : c'est à lui de "porter" son évaluation, de rendre compte de ses interprétations, de justifier la note attribuée, puisque le référent lui attribue ces marges de manœuvre.

Si nous examinons maintenant le PACT, la prétention à être fondé sur les faits (*evidence-based*) (Pecheone & Chung, 2007) nie la dimension du jugement professionnel. En effet, si l'on reprend les propos de Coles (2002), le jugement professionnel commence lorsque les limites d'un protocole sont atteintes. Les énoncés des échelles descriptives sont une forme de protocole qui normalise les attentes des évaluateurs et leur regard posé sur les pratiques observées, d'autant plus qu'ils font l'objet d'une procédure de *benchmarking* visant à faire coller l'évaluation des futurs évaluateurs, lors de leur formation, à un jugement standardisé et reconnu comme juste. Ces échelles et les commentaires qui les accompagnent visent à réduire la marge de manœuvre des évaluateurs et suspendent de fait le recours possible à leur jugement professionnel. En effet, en protocolant de manière précise le processus et le référent, ils ne permettent pas un dépassement des énoncés par l'utilisation, par exemple, d'éléments observés non répertoriés dans les descripteurs. De même, une appréciation positive qui ferait l'impasse sur un élément pourtant insuffisant, parce que globalement le candidat est perçu comme manifestant une aptitude réelle, n'est pas possible. De plus, l'impossibilité de sortir du protocole pour tenir compte du contexte du stagiaire ou des limitations dues aux contraintes de l'enregistrement vidéo ou d'éléments non nommés dans l'échelle descriptive crée une tension chez l'évaluateur, dont le coût subjectif peut être important.

Dès lors, notre analyse du PACT laisse apparaître une situation de sur-prescription, alors que celle de la HEP manifeste une sous-prescription. L'une comme l'autre de ces situations engage finalement tout autant la subjectivité de l'évaluateur. Si l'évaluation de la pratique à la HEP requiert un investissement subjectif important pour assumer la responsabilité des décisions liées à l'évaluation, le PACT, lui, opte pour une prise en charge de la plupart des décisions par le système dans le but d'obtenir une évaluation se voulant indépendante du contexte et de l'évaluateur et semble, à première vue, limiter l'investissement subjectif. Cependant, les évaluateurs se trouvent confrontés à des difficultés, lors de l'évaluation, du fait justement que le contexte de la pratique ou d'autres informations hors protocole ne sont pas pris en compte par le cadrage institutionnel. Ainsi, dans les petites universités en tout cas et selon les propos recueillis auprès des acteurs rencontrés, l'évaluateur peut être tiraillé entre sa connaissance de l'étudiant, qu'il sait apte, et le fait que cette aptitude n'est pas toujours visible sur les quelques minutes de vidéo soumises à évaluation. Rien d'étonnant ici, si l'on tient compte de la difficulté d'extraire environ vingt minutes de la pratique d'un enseignant qui soient révélatrices d'un ensemble prédéterminé d'aptitudes.

À cela s'ajoute le fait qu'un cadrage important peut engendrer une perte de sens pour l'évaluateur dans son activité professionnelle et par là de la souffrance (André, 2011, Dejours, 1993). En effet, une tension est, selon nous, perceptible entre la reconnaissance de l'expertise de cette personne par l'institution pour laquelle elle travaille comme formateur et le fait que ses actes évaluatifs soient réduits à l'identification d'éléments précis lui laissant finalement peu de marge pour faire appel à cette expertise.

6. Conclusion

Notre recherche met en évidence comment deux institutions différentes mettent en œuvre des stratégies visant à faire face aux nombreux défis que pose l'évaluation des pratiques enseignantes. Les choix effectués n'éliminent pas les difficultés, mais prennent des options sur la configuration des différents curseurs réglant les processus et les documents utilisés. En particulier, nous avons mis en lumière deux stratégies quant à la référentialisation et à l'opérationnalisation de l'évaluation de la capacité à enseigner. Les concepteurs du PACT tendent à réduire la marge de manœuvre des évaluateurs, en prenant en charge, au niveau du système, la définition de standards et descripteurs précis, sous forme d'échelles descriptives. En revanche, à la HEP, les évaluateurs ont des balises nettement moins précises, et l'essentiel de l'opérationnalisation de l'évaluation leur revient. Que ce soit par sur-prescription transformant l'évaluation en acte technique déterminé, ou par sous-prescription laissant place à l'indétermination et la subjectivité, le coût subjectif est présent dans les deux cas. Dans le premier, il est dû à la technicisation qui peut faire perdre aux évaluateurs le sens de leur activité et l'intérêt de leur expertise pour cette tâche ; dans le second, par la nécessité pour les évaluateurs de « porter » leur évaluation, dans la mesure où leurs décisions ont peu de points d'appui.

En s'intéressant au travail de l'évaluateur, cette recherche ouvre un champ important à explorer : comprendre comment se forme et se manifeste le jugement professionnel lors de l'évaluation des pratiques enseignantes. En effet, la construction de référentiels plus directement assimilables par les évaluateurs en sera sans doute facilitée. Cette compréhension de la construction des référentiels permet également de clarifier les processus pour les évalués et peut avoir un impact sur la formation des évaluateurs.

Enfin, cette étude met en évidence l'importance de construire des référentiels issus de consensus locaux et provisoires entre chercheurs, formateurs et praticiens. Ce caractère local et provisoire nous paraît le moyen d'habiter la tension entre sur- et sous-prescription, en mettant de côté tant les énoncés décontextualisés, coûteux en termes d'interprétation ou d'adaptation à la situation, qu'en effectuant un repli sur une évaluation intuitive pragmatique.

7. 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. (dans ce numéro). L'usage du référentiel de compétences par les formateurs de terrain pour évaluer les étudiants en stages.
- André, B. (2013). *S'investir dans son travail : les enjeux de l'activité enseignante*. Berne: Peter Lang.
- André, B. (2011). S'investir dans son travail: entre plaisir et souffrance. In P.-A. Doudin, D. Curchod-Ruedi, L. Lafortune & N. Lafranchise (Eds.), *La santé psychosociale des enseignants et des enseignantes* (pp. 39-59). Québec: Presses de l'Université.
- 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.
- Attali, A. & Bressoux, P. (2002). *L'évaluation des pratiques éducatives dans les premier et second degrés*. Rapport établi à la demande du Haut Conseil de l'évaluation de l'école. Repéré le 26 octobre 2015 à URL <http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/034000533.pdf>
- Berliner, D. C. (2005). The Near Impossibility of Testing for Teacher Quality. *Journal of Teacher Education*, 56(3), 205-213.
- 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., Pastré, P. & Vinatier, I. (2007). Les organisateurs de l'activité enseignante. Perspectives croisées. *Recherche et formation*, 56, 5-14.
- 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.
- 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.
- Clot, Y. (1999). *La fonction psychologique du travail*. Paris: Presses Universitaires de France.
- Coker, H., Medley, D. & Som, R. (1980). How Valid Are Expert Opinions About Effective Teaching? *Phi Delta Kappan*, 62(2), 131-134.
- Coles, C. (2002). Developing professional judgement. *The Journal of Continuing Education in the Health Professions*, 22, 3-10.
- Commission on Teacher Credentialing (2013). *California Teaching Performance Expectations*. Récupéré le 7 février 2013 du site de l'auteur.
- Crahay, M. (1989). Contraintes de situation et interaction maître-élèves. *Revue Française de Pédagogie*, 88, 67-94.
- Darling-Hammond, L., Wise, A. E. & Klein, S. P. (1999). *A License to Teach. Raising Standards for teaching*. San Francisco: Jossey-Bass Publishers.
- Davezies, P. (1993). Éléments de psychodynamique du travail. *Éducation permanente*, 116, 33-46.
- Dejours, C. (1993). *Travail, usure mentale : essai de psychopathologie du travail* (Rééd. Nouv. éd. augm. ed.). Paris: Bayard.
- Dumay, X. (2011). La recherche sur l'efficacité de l'enseignement; transformations politiques et modèle d'évaluation *Les dossiers des Sciences de l'Éducation*, 25, 57-68.
- Falzon, P. & Sauvagnac, C. (2004). Charge de travail et stress. In P. Falzon, *Ergonomie* (pp. 175-190). Paris : Presses universitaires de France.
- Figari, G. & Remaud, D. (Eds.) (2014). *Méthodologie d'évaluation en éducation et formation ou l'enquête évaluative*. Bruxelles: De Boeck.

- Figari, G. (1994). *Évaluer : quel référentiel?* Bruxelles: De Boeck.
- Gérard, F.-M. (2002). L'indispensable subjectivité de l'évaluation. *Antipodes*, 156, 26-34.
- Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- HEP. (2013a). *Master of Arts/ of Science et Diplôme d'enseignement pour le degré secondaire I*. Récupéré le 7 février 2013 du site de l'auteur.
- HEP. (2013b). *Formation des enseignants. Référentiel de compétences professionnelles*. Récupéré le 7 février 2013 du site de l'auteur.
- HEP. (2013c). *Filière Enseignement secondaire 1. Suivi du stage du semestre 4*. Récupéré le 7 février 2013 du site de l'auteur.
- InTASC. (2011). *Model Core Teaching Standards: A resource for State Dialog*. CCSSO's Interstate Teacher Assessment and Support Consortium
- Jobert, G. (1999). L'intelligence au travail. In P. Carré & P. Caspar (Eds.), *Traité des sciences et des techniques de la formation*. Paris: Dunod.
- Lenoir, Y. (2010). La notion de référentialité dans la formation à l'enseignement : apports et limites. *Recherche et formation*, 64, 91-104.
- Leplat, J. (1997). *Regards sur l'activité en situation de travail : contribution à la psychologie ergonomique*. Paris: Presses universitaires de France.
- Louis, R., Jutras, F. & Hensler, H. (1996). Des objectifs aux compétences: implications pour l'évaluation de la formation initiale des maîtres. *Canadian Journal of Education / Revue canadienne de l'éducation*, 21(4), 414-432.
- McDonald, F. J. & Elias, P. (1976). *Executive Summary Report : Beginning Teacher Evaluation Study, Phase II*. Princeton: University of Princeton, Educational Testing Service.
- Mitchell, K. J., Robinson, D. Z., Plake, B. S. & Knowles, K. T. (Eds.). (2001). *Testing Teacher Candidates. The Role of Licensure Tests in Improving Teacher Quality*. Washington, DC: National Academy Press.
- 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.
- Ombredane, A. & Faverge, J.-M. (1955). *L'analyse du travail : facteur d'économie humaine et de productivité*. Paris: Presses Universitaires de France.
- PACT Consortium. (2010). *Elementary Mathematics Rubrics*. Last updated 2012
- Pastré, P., Mayen, P. & Vergnaud, G. (2006). La didactique professionnelle. *Revue française de pédagogie*, 154, 145-198.
- Pecheone, R. L. & Chung, R. R. (2007). *Technical Report of the Performance Assessment for California Teachers (PACT): Summary of Validity and Reliability Studies for the 2003-04 Pilot Year*. Stanford University: PACT Consortium.
- Périsset, D. & Bagnoux, D. (2007). Les référentiels de compétences, profils attendus de la professionnalité enseignante. Acteurs et points de vue en Suisse romande. In L. Talbot & M. Bru (Eds.), *Des compétences pour enseigner* (pp. 87-104). Rennes: Presses universitaires.
- 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.
- Perrenoud, P. (2001). Evaluation formative et évaluation certificative : postures contradictoires ou complémentaires ? *Formation professionnelle suisse*, 4, 25-28.
- Peterson, K., & Kauchak, D. (1982). *Teacher Evaluation: Perspectives, Practices, and Promises*. Salt Lake City: Center for Educational Practice, University of Utah.
- Prost, A. (1986). *Éloge des pédagogues*. Paris: Seuil.
- Roegiers X. (2000) *Une pédagogie de l'intégration : compétences et intégration des acquis dans l'enseignement*. Bruxelles : De Boeck.

- Royce Sadler, D. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, 34(2), 159-179.
- Scallon G. (2004). *L'évaluation des apprentissages dans une approche par compétence*. Québec : Éditions du Renouveau pédagogique.
- 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.
- Tupin, F. (coord.) (2003). De l'efficacité des pratiques enseignantes ? *Les Dossiers des Sciences de l'Éducation*, 10, Toulouse : PUM.
- Vinatier, I. (2008). *Analyser et comprendre la pratique enseignante*. Rennes: Presses Univ. de Rennes.
- Wegener Soled, S. (Ed.). (1995). *Assessment, testing and evaluation in teacher education*. Norwood: Ablex Publishing Corporation.
- Weiss, J. (1986). La subjectivité blanchie? In J.-M. De Ketele (Ed.), *L'évaluation : approche descriptive ou prescriptive?* (pp. 91-105). Bruxelles: De Boeck.

