

Analyse des outils d'accompagnement et d'évaluation : point de vue des formateurs de terrain

Bélair Louise

Université du Québec à Trois-Rivières

Louise.Belair@uqtr.ca

Lebel Christine

Université du Québec à Trois-Rivières

Christine.Lebel@uqtr.ca

Monfette Olivia

Université du Québec à Trois-Rivières

Olivia.Monfette1@uqtr.ca

Miron Geneviève

Commission scolaire des Chênes, Québec

genemiron@hotmail.com

Blanchette Sylvie

Commission scolaire des Chênes, Québec

bsy@cgocable.ca

Résumé

Le groupe de recherche collaborative REÉVES poursuit depuis 3 ans son étude des pratiques de jugement d'évaluation des enseignants associés (formateurs de terrain) par le biais de différents chantiers de travail. Cet article, issu de l'un des chantiers, a pour but de donner la parole aux formateurs de terrain (FT) pour analyser l'efficacité, l'efficacé et la pertinence des outils d'évaluation et des échelles descriptives sur leur prise de décision quant au niveau de compétence développé par le stagiaire. À partir des guides de stage élaborés par différentes universités québécoises, l'équipe a répertorié divers outils afin que des formateurs de terrain puissent, dans le cadre de groupes de discussion et d'entretiens, expliquer en quoi tel outil est plus facilitant, pertinent et utile dans leur prise de décision (Figari et Remaud, 2014 ; Tricot et Tricot, 2000). Ces outils sous forme de grilles sont souvent conçus par les responsables des programmes concernés, parfois en collaboration avec des superviseurs universitaires et plus rarement avec des formateurs de terrain (Bélair, Vivegnis et Lafrance, 2015 ; Lapointe et Guillemette, 2015 ; Portelance, 2010). Ils leur servent de cadre de référence car ils sont le reflet des orientations de ces programmes qui, malgré un référentiel commun, diffèrent passablement et ce, parfois au sein d'une même université. Les formateurs de terrain doivent ainsi constamment s'ajuster à des orientations différentes et des formes d'outils fort diversifiés, ce qui semble engendrer des tensions dans la manière de compléter ces grilles et de prendre une décision. Les premiers résultats de cette recherche évaluative montrent sans équivoque qu'aucune de ces grilles ne répond parfaitement à leurs besoins. Toutefois une analyse des verbatim permet de dégager ce qui serait utile, cohérent et facilitant pour prendre une décision éclairée et équitable envers le stagiaire. Cet article présente des liens potentiels entre ces choix de grilles et d'échelles et l'impact sur la décision du niveau de compétence développé et ouvre sur des pistes futures d'investigation.

Mots-clés

Outils pour accompagner et évaluer des stagiaires ; formateurs de terrain ; stages.

Abstract:

Over the past three years, the RÉEVES Collaborative Research Group has studied the evaluation practices of cooperating teachers (CTs) through various field projects. The present article, based on one of these projects, seeks to gain input from CTs and to analyze the efficiency, effectiveness and pertinence of assessment tools and descriptive scales with regard to how CTs make decisions about the proficiency levels of student teachers. Drawing on evaluation grids from practicum handbooks developed by universities in the province of Quebec, the team asked CTs to identify which ones are the most helpful, pertinent and useful for decision making when evaluating (Figari & Remaud, 2014; Tricot & Tricot, 2000). These grids are typically developed by program heads, sometimes in cooperation with university supervisors and, more rarely, with cooperating teachers (Bélair, Vivegnis & Lafrance, 2015; Lapointe & Guillemette, 2015; Portelance, 2010). They also provide a reference framework because they reflect program focus, which often differs considerably from one program to another—even within the same university. However, CTs must constantly adapt to different program objectives and evaluation tools, causing unease and confusion over how to fill out grids and make evaluation decisions. Preliminary results suggest that none of the proposed grids perfectly meet their needs. Verbatim analysis from focus group helped identify useful, coherent and facilitating elements for making informed and fair decisions about student teachers. Possible links between the choice of grids and scales, and the potential impact on evaluating proficiency level is also discussed and avenues for further research are proposed.

Pour citer cet article : Bélair, L.M., Lebel, C., Monfette, O., Miron, G. & Blanchette, S. (2016). Analyse des outils d'accompagnement et d'évaluation : point de vue des formateurs de terrain. *Evaluer. Journal international de Recherche en Education et Formation*, 2(2), pp. 9-25.

1. Introduction

Le groupe de recherche collaborative RÉEVES¹ de l'UQTR mène depuis 3 ans une étude sur les pratiques de jugement d'évaluation d'enseignants associés (formateurs de terrain²). À la suite de la mise sur pied de différents chantiers permettant de référentialiser ces pratiques (Lebel, Bélair et Monfette, 2016 ; Lebel, Bélair, Monfette, Hurtel, Miron et Blanchette, 2016), l'un d'entre eux a porté sur l'analyse de l'outillage susceptible d'aider le formateur de terrain à accompagner et à évaluer le stagiaire afin de prendre une décision éclairée et équitable. Ce présent article propose ainsi les résultats de cette recherche évaluative.

2. Problématique

Au Québec, la formation à l'enseignement est dispensée par les universités dans le cadre de programmes de formation initiale professionnalisants. Ces programmes doivent tenir compte des prescriptions ministérielles qui préconisent le développement de 12 compétences professionnelles par les stagiaires dans le cadre d'une formation de quatre ans, composée d'un ensemble de cours et de 700 heures de stages. Le CAPFE³ s'assure de l'adéquation entre ces divers programmes et orientations prises localement par les universités et les attentes ministérielles. En ce qui concerne les stages, les institutions de formation fournissent des guides et des outils d'accompagnement et d'évaluation et offrent à des degrés divers, des

¹ RÉEVES signifie le Regroupement pour l'Étude de l'Évaluation des Enseignants Stagiaires. <http://www.regrouperreeves.ca>

² Comme les appellations divergent d'un pays à l'autre pour désigner le maître de stage, l'enseignant associé, l'enseignant hôte, etc., nous avons convenu d'utiliser un terme générique qui englobe toutes ces appellations, soit : formateur de terrain.

³ CAPFE : Comité d'agrément des programmes de formation à l'enseignement.

formations pour les formateurs de terrain basées sur un cadre de référence pour l'accompagnement du stagiaire.

Plusieurs études québécoises font état du processus d'élaboration des guides de stage par les universités et des modifications qui y ont été apportées depuis l'avènement du référentiel de compétences en 2001 (Brochu et Bouvier, 2008 ; Gagnon et Mazalon, 2008 ; Gauthier et Mellouki, 2006 ; Ouellet, Pache et Portelance, 2008). Elles montrent le virage qu'a connu l'évaluation, qui met en avant l'importance de témoigner de la progression des compétences professionnelles des étudiants, notamment lors des stages (Ouellet et al., 2008). Cette orientation a entraîné une refonte des guides de stages et des outils d'évaluation afin qu'ils permettent effectivement de porter un jugement sur le niveau de développement des compétences professionnelles des stagiaires. Toutefois, bien qu'il soit possible de percevoir les efforts mobilisés pour cadrer avec l'approche par compétences prescrite par le MÉLS⁴, ces écrits permettent de soulever deux constats.

2.1 Premier constat : la divergence des outils offerts aux formateurs de terrain

On constate en premier lieu que les guides de stage divergent d'un programme et d'une université à l'autre. Ces outils servant de cadre de référence pour les formateurs de terrain reflètent les orientations des programmes desquels ils sont issus mais, en dépit d'un référentiel de 12 compétences professionnelles unique et prescrit, ils prennent des formes très contrastées et ce, parfois au sein d'une même université (Camaraire et Cournoyer, 2002). Chaque guide de stage est ainsi teinté d'une conception particulière de l'évaluation, ce qui entraîne en conséquence des choix divergents en ce qui concerne les niveaux de compétences et les seuils de réussite (Bélaïr, 2007 ; Boyer, 2013). Or, comme les formateurs de terrain sont souvent confrontés au cours d'une année scolaire à plusieurs outils provenant de différentes universités, cela nécessite une forme d'ajustement de leur part, ce qui peut causer des tensions dans la manière de compléter ces grilles et de prendre une décision éclairée.

Outre la nécessité de s'adapter aux différents guides de stage, les formateurs de terrain se retrouvent également seuls pour se les approprier et ce, malgré quelques courtes formations ou documents explicatifs, ce qui a pour effet que chacun d'eux construit sa propre conception des compétences professionnelles et des grilles d'évaluation (Bates et Burbank, 2008; Desbiens, Spallanzani, et Borges, 2013 ; Gouin et Hamel, 2015). Ceci porte à croire que l'évaluation des stagiaires est d'autant plus teintée de subjectivité que chaque formateur de terrain interprète de manière arbitraire les grilles d'évaluation fournies par les universités, ce qui peut nuire à la validité et à l'équité du processus (Tardif, 2006 ; Scallon, 2004).

2.2 Deuxième constat : le manque d'implication des formateurs de terrain dans l'élaboration des outils d'accompagnement et d'évaluation

Le deuxième constat s'impose d'emblée puisque les guides de stages sont rarement élaborés en concertation avec les formateurs de terrain. Les formateurs de terrain sont peu impliqués dans le processus d'élaboration des guides et des critères d'évaluation (Lebel et Bélaïr, sondage informel mené auprès du groupe REÉVES, communication inédite, 2013), conçus pour les aider à accompagner et à évaluer les stagiaires (Gagnon et Mazalon, 2008). Une enquête menée à l'UQTR, (Lapointe et Guillemette, 2015) montre pourtant la nécessaire participation de ces acteurs de première ligne.

⁴ Bien que le MÉLS se nomme depuis septembre 2015 le MEESR : Ministère de l'éducation, de l'enseignement supérieur et de la recherche, nous utilisons l'acronyme qui était en vigueur au moment de la recherche.

Ce manque apparent de concertation pourrait alors expliquer pourquoi plusieurs recherches mettent en avant le fait que les formateurs de terrain mentionnent que les guides et outils semblent élaborés sans égard au contexte professionnel (Bélaïr, Vivegnis et LaFrance, 2015 ; Lapointe et Guillemette, 2015 ; Portelance et Tremblay, 2006 ; Portelance, 2010). Plusieurs formateurs les conçoivent comme des documents hermétiques et difficiles d'utilisation, voire même parfois inutiles, ce qui contribue à rendre le processus d'évaluation des stagiaires laborieux pour eux (Gervais et Lepage, 2000 ; Gouin et Hamel, 2015 ; Miron et Blanchette, 2015). On constate ainsi que de tels outils prescrits par les universités, même s'ils sont accompagnés de formations ou de documents explicatifs, ne revêtent pas la même légitimité que des outils construits en collaboration et en situation. Cette légitimité institutionnelle reconnue par les membres de la communauté universitaire n'est pas garante d'une légitimité locale et vécue par le formateur de terrain aux prises avec la réalité quotidienne de sa classe (Figari et Remaud, 2014). En conséquence, cette situation peut faire en sorte que les formateurs de terrain se sentent relégués au rang de simples utilisateurs et qu'ils envisagent ces outils comme étant normatifs, sans vision partagée avec les partenaires universitaires responsables de l'évaluation en stage (*ibid.*, 2014).

Les rares résultats de recherche à ce sujet (Gouin et Hamel, 2015 ; Lapointe et Guillemette, 2015 ; Monnard et Luisoni, 2013) préconisent de mener une réflexion plus approfondie sur le dispositif général et la portée de ces outils, d'autant plus qu'ils semblent presque essentiellement perçus par les formateurs de terrain comme des référentiels aux fins d'évaluation certificative des compétences professionnelles des stagiaires (Lapointe et Guillemette, 2015). Ceci permettrait alors de poser un regard sur les enjeux impliqués dans une prise de décision équitable du formateur de terrain envers le stagiaire. De fait, le sens accordé à l'évaluation des stagiaires par les formateurs de terrain est en partie teinté par les outils fournis par les universités, mais ce sens est aussi imprégné de l'appartenance idéologique, de l'épistémologie dominante qui supporte la conception de ces outils et de l'implicite de l'évaluateur et de la situation (Figari et Remaud, 2014). Dans un tel contexte, on peut supposer que le formateur de terrain se fait une idée tacite de ce que devrait faire un stagiaire à partir des outils fournis à même le guide de stage, de ceux déjà utilisés et qui proviennent d'autres universités et finalement, à partir de sa propre conception de ce qu'est l'enseignement. Il semble ainsi pertinent d'envisager que l'évaluation des stagiaires est une source de questionnements et d'inquiétudes pour les formateurs de terrain (Lapointe et Guillemette, 2015).

Au vu de tels constats, le but de la présente recherche est de donner la parole à des formateurs de terrain pour analyser des outils d'accompagnement et d'évaluation au regard d'une prise de décision équitable quant aux niveaux de compétences développées par le stagiaire. La question de recherche se pose comme suit : selon des formateurs de terrain, quelles sont les caractéristiques d'un outil d'accompagnement et d'évaluation susceptible de les aider à prendre une décision équitable dans l'évaluation des compétences des stagiaires ?

3. Cadre de référence

3.1 Des outils d'évaluation et d'accompagnement, à l'aune des conceptions de l'évaluation

Le jugement professionnel du formateur de terrain exige de la rigueur et de la planification puisqu'il doit analyser les défis à relever par le stagiaire et le soutenir dans l'ensemble du développement de ses compétences. L'évaluation des stagiaires implique des jugements holistiques qui englobent les conceptions personnelles de la qualité de l'acte par l'évaluateur,

l'interprétation de la performance des stagiaires et des critères d'évaluation fournis par l'établissement d'études, de même que la prise en considération du contexte et du milieu de stage (Tang, 2008 ; Knight, 2006 ; Leshem et Bar-Hama, 2008). Pour y arriver, le formateur doit donc décoder par le biais d'observations pour ensuite rétroagir auprès de ce stagiaire (Leroux et Bélair, 2016), en s'appuyant sur différents types de grilles d'évaluation prescrites : des échelles uniformes qui regroupent des termes qualitatifs, des symboles ou des nombres (par exemple les symboles + ou -), des échelles descriptives qui présentent des attentes (niveaux décrits par compétence) et des échelles descriptives globales qui indiquent des critères d'évaluation plus généraux pour chaque compétence ciblée (atteint / non atteint) (Gouin et Hamel, 2015 ; Scallon, 2004). Les moments de régulation, partie prenante de l'accompagnement dans un processus continu d'interactions associées à l'évaluation formative (Leroux, 2014), permettent ainsi au formateur de terrain de poser des diagnostics et de commenter les planifications ou les enseignements du stagiaire à partir des grilles et des outils fournis dans les guides (Allal et Mottier Lopez, 2007).

Le processus d'évaluation certificative du stagiaire se fait à partir de jugements menant à une décision d'ordre administratif liée à la note et à la réussite du stage, même si le jugement relève aussi de l'implicite et de la situation, par exemple lorsque les formateurs de terrain « ne savent comment dire à un étudiant qu'il est en échec sans le blesser et pouvoir poursuivre la collaboration dans de bonnes conditions » (Van Nieuwenhoven et Colognesi, 2013, p. 134). Ce processus se réalise certes à la suite des traces cumulées tout au long du stage, impliquant la mise en relation de différentes informations (observations, discussions, notes, etc.) pour établir le niveau de compétence développé par le stagiaire (De Ketele, 2010). Cependant, on constate que, dans les guides de stage, ce jugement se traduit différemment et reflète plusieurs conceptions de l'évaluation, allant des cotes S (succès) et É (échec) à une échelle de cotes associée à des descripteurs quantitatifs (par exemple : parfois, toujours, etc.) ou qualitatifs (par exemple : correspond au niveau de maîtrise attendue ; correspond partiellement, etc.) ou encore à des notes chiffrées (Bélair, 2007 ; Brochu et Bouvier, 2008 ; Ouellet et al., 2008). Une réflexion entourant la scientificité de ces différents outils s'avère donc fort pertinente.

3.2 Vers une recherche évaluative

Dans le but de donner la parole à des formateurs de terrain pour analyser des outils d'accompagnement et d'évaluation au regard d'une prise de décision équitable quant aux niveaux de compétences développées par le stagiaire, cette recherche évaluative adopte une approche qui se rapproche du « Client-Oriented Model » (Scriven, 1967 ; Stufflebeam, 2001 ; Vedung, 2005). Elle facilite l'évaluation de l'efficacité d'un outil en questionnant les utilisateurs de l'outil et non pas ses développeurs et permet l'identification des attentes, des préoccupations et des besoins des utilisateurs (Belnat, Leménager et Mermet, 2012). Pour ce faire, cette étude se base sur des critères issus de diverses recherches évaluatives, notamment dans le domaine de l'informatique s'appuyant sur les travaux de Tricot et Tricot (2000), dans le domaine des politiques environnementales s'appuyant sur les travaux de Belnat et al. (2012) ainsi que dans le domaine de l'évaluation de programme à partir des travaux de Figari et Remaud (2014). Les critères d'efficacité, de pertinence et d'efficacités sont définis et servent d'ancrage théorique.

3.3 Efficience

L'analyse des outils par les formateurs de terrain doit permettre, dans un premier temps, de soustraire des informations relatives à leur utilisabilité et à leur utilité, au sens de la pertinence situationnelle tel que proposé par Simonnot (2008). Ainsi que l'indiquent Tricot et Tricot (2000) : « l'utilisabilité d'un objet finalisé désigne la possibilité d'utilisation de cet objet [alors que] l'utilité (...) désigne la possibilité d'atteindre le but visé avec cet objet » (p. 3). En ce sens, l'utilisabilité s'apparente à la facilité d'utilisation des outils par les formateurs de terrain. Pour sa part, le critère d'utilité, selon le principe d'une approche descendante de type *bottom down*, vise à vérifier si cet outillage permet aux formateurs de terrain d'atteindre les buts pour lesquels ils ont été initialement conçus, soit de permettre facilement l'évaluation des compétences professionnelles des stagiaires (Bélaïr et al., 2015 ; Tricot et Tricot, 2000). Ceci rejoint la notion d'efficience au sens de Figari et Remaud (2014) qui mentionnent que l'efficience d'un dispositif se joue sur la relation entre les enjeux et les résultats, la rentabilité des investissements et les techniques proposées. Il va donc de soi que l'efficience des outils présentés aux participants se manifeste par la relation entre le point de vue externe du concepteur et celui de l'utilisateur. Dans le cadre de cette recherche, l'efficience des outils se présente comme la manière dont les outils sont utiles et utilisables pour les utilisateurs en fonction de leur simplicité et de l'économie de temps.

3.4 Pertinence cognitive

Dans une lignée fort similaire à l'efficience, la pertinence représente l'adéquation entre la finalité de l'outil et le but de l'utilisateur (Figari et Remaud, 2014, Mizzaro, 1997). Ce critère, selon le principe d'une approche ascendante de type *bottom up*, permet, cette fois, de mettre en relation les données de contexte de stage avec les choix adoptés par les concepteurs des outils et ce, afin de valider les choix au vu d'une décision juste et équitable. Ainsi, pour que les outils soient pertinents, ils doivent effectivement permettre aux formateurs de terrain de procéder à l'évaluation des compétences professionnelles des stagiaires. Il s'agit donc de s'assurer que le besoin tel qu'exprimé par l'utilisateur est comblé par les outils proposés, en termes de vocabulaire, de concepts et de facilité de compréhension et non pas de cibler la pertinence au sens de la validité des outils présentés. Le tout en vue d'analyser l'adéquation entre ce qui est attendu et compris des formateurs de terrain et ce qui est proposé dans les documents. Dans cette recherche, la notion de pertinence cognitive est ainsi reliée à un besoin d'informations de la part de l'utilisateur qui exprime ou cherche des informations en lien avec son besoin qui est souvent exprimé dans un langage usuel (Simonnot, 2008).

3.5 Efficacité

L'efficacité d'un outil consiste en la mise en relation entre ce qui est attendu et ce qui est obtenu (Figari et Remaud, 2014) par l'analyse des observations sur le terrain en lien avec les prescriptions des programmes présentés par les outils. Pour Belnat et al. (2012), un outil efficace est celui qui permet d'atteindre les objectifs fixés par les dispositifs prescrits, ce qui rejoint quelque peu le critère d'utilité de Tricot et Tricot (2000), sans toutefois entrer dans les enjeux d'économie de temps, mais en mettant plutôt l'accent sur le fait que l'outil permet au formateur de terrain de parvenir à ses fins, soit d'évaluer le stagiaire de manière équitable. En s'appuyant sur ces définitions, dans le cadre de nos propres travaux, l'efficacité des outils repose sur la possibilité d'atteindre le but recherché, soit de porter un jugement d'évaluation équitable au regard du développement des compétences professionnelles des stagiaires.

4. Méthodologie

Le recueil des données a été réalisé dans le cadre de groupes de discussion et d'entretiens à partir d'un questionnaire⁵ composé de 12 questions demandant aux participants d'exercer un choix entre différents documents en fonction de leur clarté, leur utilité, le nombre d'indicateurs, etc. et ce, en lien avec les critères d'efficacité, de pertinence cognitive et d'efficacité. Dix-huit documents dont des extraits anonymes de guides de stages, des modèles de grilles d'observation, de grilles d'accompagnement et de régulation et d'échelles descriptives certificatives ont été répertoriés à partir des guides fournis par plusieurs programmes de formation à l'enseignement de quatre universités québécoises. Ces documents concernent principalement les compétences professionnelles relatives à la conception de situations enseignement apprentissage et au pilotage (CP3 et CP4)⁶.

Pour chacun des éléments étudiés, les participants devaient choisir individuellement entre deux ou trois modèles de documents qui leur étaient présentés et expliquer la nature de leur choix (individuellement ou de manière collective). Les deux enseignantes participant au projet REÉVES ont ainsi mené des groupes de discussions et des entretiens auprès de 41 formateurs de terrain dont 28 au primaire et 13 du secondaire, issus de quatre commissions scolaires et ayant déjà suivi la formation des enseignants associés offerte par les différentes universités québécoises.

Cette recherche évaluative de nature qualitative visait à donner la parole aux formateurs de terrain et ainsi cherchait à identifier quelles sont les caractéristiques d'un outil facilitant un jugement d'évaluation équitable envers un stagiaire. Ainsi, le fait d'avoir demandé de choisir entre des documents en fonction de certaines caractéristiques a servi d'ancrage pour les discussions qui s'en sont suivies. Les choix ont certes été analysés en fonction du pourcentage de réponses, mais l'ensemble des analyses a reposé sur les verbatim issus des discussions lors des entretiens et des groupes de discussion, à partir d'une analyse de contenu manuelle effectuée par le groupe de recherche.

5. Résultats

Les résultats reflètent l'analyse des choix effectués et expliqués par des formateurs de terrain en fonction des critères visés, soit l'efficacité, la pertinence cognitive et l'efficacité. Ils sont généralement discutés sans égard à l'ordre d'enseignement des répondants, sauf dans quelques cas où il y a des écarts significatifs entre les réponses des enseignants du primaire et ceux du secondaire. Afin d'explicitier ces résultats et d'entrée de jeu les discuter, il est apparu intéressant de combiner leur présentation justifiée par des verbatim et leur analyse appuyée par des recherches similaires dans le domaine.

⁵ Le questionnaire et la recherche de documents ont été réalisés par Sylvie Blanchette et Geneviève Miron de la Commission scolaire Des Chênes.

⁶ Compétence CP3 : Concevoir des situations d'enseignement-apprentissage pour les contenus à faire apprendre, et ce, en fonction des élèves concernés et du développement des compétences visées dans le programme de formation.

Compétence CP4 : Piloter des situations d'enseignement-apprentissage pour les contenus à faire apprendre, et ce, en fonction des élèves concernés et du développement des compétences visées dans le programme de formation.

5.1 L'efficience

Ce critère vise la relation entre le point de vue externe du concepteur et celui de l'utilisateur dans la manière de voir si les outils sont utilisables, utiles et économes en temps. Pour ce faire, les outils ont été analysés sous quatre angles : la quantité d'indicateurs proposés dans une grille, la facilité d'utilisation d'une grille, l'utilité d'autres documents annexés aux guides et l'utilité de diverses grilles adressées aux autres partenaires de la triade.

Pour connaître la préférence des participants quant à la **quantité d'indicateurs** nécessaire pour fins de régulation, deux grilles d'évaluation de la compétence CP3 ont été proposées. Ces deux modèles de grilles ont été choisis en raison de leur grande différence dans l'orientation de la régulation à offrir aux stagiaires. De fait, l'une est une liste de 38 comportements observables à cocher, tandis que l'autre présente plutôt huit indicateurs de comportements observables et une section permettant des commentaires généraux.

Les résultats indiquent que les formateurs de terrain préfèrent à 58% la grille qui contient huit indicateurs et un espace prévu pour des commentaires généraux. Toutefois, on constate que 40% des participants, surtout du secondaire, préfèrent la grille de 38 comportements à cocher. Ceci confirme en partie les résultats de Lapointe et Guillemette (2015) en ce qui a trait au nombre d'indicateurs parfois perçu comme trop grand chez les participants. Toutefois, les propos d'un répondant permettent de percevoir l'utilité des indicateurs dans certains cas précis :

- *Une grille ayant de nombreux indicateurs de comportements observables m'apparaît intéressante dans les moments où on a à poser un regard plus pointu, à se faire une tête par rapport à une stagiaire en difficulté (FS02⁷).*

Au regard de la **facilité d'utilisation d'une grille**, les participants devaient choisir entre deux modèles de grilles pour consigner leurs observations sur la compétence CP3. La première présente huit indicateurs de comportements observables suivis d'un espace permettant des commentaires généraux, la seconde présente toutes les composantes de la compétence suivies d'un carré à cocher, sans espace de commentaires. Ainsi, dans la première grille, les indicateurs servent d'ancrage au commentaire tandis que dans la deuxième, les concepteurs dirigent la réflexion sur des composantes détaillées de la compétence.

Tel que repéré au niveau de la préférence de la quantité d'indicateurs, une bonne proportion des enseignants (65%) tant du primaire que du secondaire préfère des grilles d'observation qui, tout en ayant quelques indicateurs, laissent place aux commentaires (grille #1). Ils veulent écrire leurs propres commentaires car, selon eux, cela leur permet de nuancer l'évaluation du stagiaire et de lui apporter une aide ponctuelle et contextualisée. Dans cette optique, des participants perçoivent les grilles comme des outils « idéels » (Figari, 2013) conçus au regard de modèles et de valeurs peu en lien avec les contextes particuliers des stages. Malgré tout, environ 35% des enseignants choisissent le deuxième outil qui présente une liste de composantes à cocher. Cette grille, plus rapide à remplir, leur permet d'avoir une vision du nombre de crochets en un seul coup d'œil et facilite la décision. Ce choix n'est pas anodin. Au moment de l'observation *in situ*, il est très difficile de prendre des notes sans avoir pu prendre le temps de bien connaître l'objet d'observation (Bélaïr, 2014) puisque tout observateur doit planifier son observation et la cadrer à l'aide d'un outil précis, lui évitant ainsi de suivre le mouvement, de noter de manière aléatoire et finalement, de s'écarter de l'objet à observer (Bélaïr, 2005). Le fait de cocher une liste d'éléments permet ainsi de passer

⁷ F ou M réfèrent à un homme ou une femme. P ou S réfèrent à un enseignant du primaire ou du secondaire.

outre la planification de l'observation pour suivre le mouvement créé par le concepteur de la grille.

Pour poursuivre cette réflexion, il est apparu intéressant de savoir si des **documents autres que ceux prévus** explicitement pour l'évaluation étaient utiles pour effectuer un jugement d'évaluation. Afin de faciliter la discussion, les participants pouvaient consulter trois documents complémentaires. Il s'agit d'un questionnaire pour favoriser la réflexion du stagiaire sur sa pratique, d'un exemple d'une journée typique de stage et d'un canevas d'une planification détaillée.

Tant au primaire qu'au secondaire le document pour guider la réflexion du stagiaire et celui concernant une planification détaillée sont appréciés par la moitié des participants. Les extraits qui suivent montrent bien la nuance que certains apportent quant à l'utilité de tels documents dans un guide de stage.

- *Plus utiles lors d'une première expérience comme enseignant associé (FP12).*
- *Ça m'aide à bien comprendre le rôle que l'université voudrait que je joue (MS22).*

En ce qui concerne le document présentant la journée typique de stage, il est jugé utile par tous les enseignants du secondaire mais il ne rejoint que le 1/3 des enseignants du primaire. L'écart semble s'expliquer par la manière dont les différents milieux scolaires vivent leur journée de classe. De fait au primaire, la classe est un milieu où il s'agit d'intégrer le stagiaire à la vie d'un seul groupe, ce qui rend le document moins nécessaire. Toutefois au secondaire, le milieu étant moins centré sur la vie quotidienne de l'élève et le nombre de groupes classes étant plus grand qu'au primaire, les enseignants doivent gérer leur temps entre les groupes classes et ressentent le besoin de se coordonner, d'où l'intérêt pour un document montrant une journée typique de stage pour un stagiaire au prise avec cette réalité professionnelle.

Enfin, les participants se sont prononcés sur **l'utilité des différentes grilles** fournies dans les guides, notamment celles à remplir par le superviseur et par le stagiaire lui-même. Le tiers des enseignants tant du primaire que du secondaire disent ne pas voir l'utilité de connaître les outils que doivent remplir les autres membres de la triade. Ceci laisse ainsi supposer que la collaboration recherchée et appréciée tant par les formateurs universitaires que de terrain lors de recherches antérieures (Gervais et Lepage, 2000 ; Lapointe et Guillemette, 2015 ; Portelance et Tremblay, 2006), semble s'atténuer au moment du stage puisque les formateurs de terrain ne ressentent pas le besoin de consulter les outils qui s'adressent aux autres partenaires de la triade.

- *Je me réfère seulement à ce qui me concerne (FS41).*

En ce qui a trait à l'efficacité, on remarque que les propos des participants sont très nuancés. On peut, par contre, noter que les formateurs de terrain semblent préférer ne pas avoir trop d'indicateurs afin que l'accent soit mis sur leurs commentaires personnalisés, qu'ils apprécient que certains documents s'adressant au stagiaire (réflexion, planification) soient intégrés au guide et qu'ils voient moins l'utilité d'inclure les outils d'évaluation des superviseurs.

5.2 La pertinence cognitive

Ce critère se fonde sur un seul élément, soit **la pertinence des indicateurs pour évaluer une compétence telle que perçue par les répondants**. En lien avec l'efficacité, il permet de poser un regard plus approfondi sur l'adéquation des indicateurs en fonction des situations particulières vécues dans les stages et des implicites théoriques sous-jacents à ces outils. De fait, les répondants devaient partir de leurs vécus et de leur compréhension des compétences pour identifier les documents qui leur semblaient les plus pertinents. Afin de susciter la

réflexion, trois grilles d'évaluation pour la compétence CP3 leur ont été présentées. La première est une liste de huit indicateurs à cocher, la deuxième présente cinq objectifs spécifiques contenant chacun entre deux et six indicateurs pondérés ainsi qu'un espace pour des commentaires à chaque objectif, et la troisième grille définit cinq indicateurs à noter individuellement et deux espaces pour des commentaires sous forme de points forts et points à améliorer. On peut ainsi associer ces trois modèles de grilles à des conceptions épistémologiques implicites ou explicites chez les concepteurs, ce qui permet en quelque sorte de dégager les tendances que semblent privilégier les différents programmes. Cette nécessaire analyse de la pertinence présuppose que, pour reprendre les propos de Figari (2013), la théorisation de l'évaluation apparaît dans un carrefour de disciplines sociales dont l'élément catalyseur est la mesure (quantitative ou qualitative) et qui sous-tend une configuration de paradigmes construits tout au long des réflexions des concepteurs.

La première grille reflète la notion de liste de vérification très prisée lors des observations (Bélaïr, 2014). Elle favorise une forme de transfert entre le moment de l'observation et celui du jugement et risque ainsi de susciter une plus grande cohérence entre ces deux temps de l'évaluation par le formateur de terrain.

La deuxième grille est d'inspiration behavioriste car elle fait l'hypothèse que la compétence se détaille en une série d'objectifs à atteindre et qu'une pondération quantitative répond du développement de ladite compétence. Dans un tel cas, on peut supposer un glissement d'une approche privilégiée dans une approche par compétences, soit le socioconstructivisme, vers un retour à un certain confort dans le connu de l'évaluation, soit une approche par objectifs (Nguyen et Blais, 2007).

Finalement, la troisième grille semble répondre aux exigences d'une approche par compétences puisqu'elle présente à la fois des indicateurs et un espace pour des commentaires centrés sur la régulation des apprentissages du stagiaire (Leroux et Bélaïr, 2015).

La presque totalité des enseignants sondés (90%), tant au primaire qu'au secondaire, disent vouloir des grilles qui leur donnent la possibilité de faire des commentaires tout en leur fournissant une liste de comportements facilement observables qui décrivent de façon claire et précise les attentes envers le stagiaire (soit les grilles 2 et 3). Une bonne partie d'entre eux (40%) apprécie la possibilité de cocher individuellement les indicateurs (grilles 1 et 3). Ceci peut s'avérer logique dans un contexte où les programmes sont centrés sur des compétences à développer, car ces participants ont opté pour des modèles de grilles qui semblent avoir été conçus en harmonie avec l'approche privilégiée par le MÉLS.

- *J'apprécie que ce soit très complet et détaillé. Ça reflète bien ce que le stagiaire a à faire en classe. Les critères s'y retrouvent et il y a un bon mélange entre les commentaires et les indicateurs (FP08).*

Toutefois, on constate que 60% des enseignants du secondaire portent leur choix sur une grille ayant des objectifs spécifiques et des comportements observables détaillés et pondérés. On peut aussi penser que le fait que cette grille comportait des résultats chiffrés y a été pour beaucoup dans leur préférence. Cette tendance fortement inspirée de leurs méthodes d'évaluation envers leurs élèves n'est pas surprenante puisque Perrenoud (2001) soutient que le formateur de terrain part de ses propres gestes pour regarder et apprécier l'Autre et qu'il est évidemment plus confortable pour lui d'utiliser une échelle connue pour évaluer les compétences professionnelles des stagiaires. Des études (Bélaïr et Dionne, 2010 ; Bélaïr et Baribeau, 2010) notent à cet égard que la multiplicité des réformes en éducation au Québec a fait en sorte que plusieurs enseignants se sont repliés sur des techniques connues et confortables pour eux plutôt que de tenter l'expérience de nouveaux défis peu rassurants.

Au regard de la pertinence cognitive et à la lumière de ce qui précède, force est d'admettre que les pratiques d'évaluation dans la salle de classe influencent les choix des formateurs de terrain. Aussi, on remarque que les formateurs de terrain ne semblent pas faire de lien entre le contexte et les conditions inédites à chaque stage et les indicateurs présentés dans les grilles, d'où leur besoin d'écrire des commentaires. Cette situation peut probablement s'expliquer en bonne partie par le fait qu'ils voient ces outils comme étant normatifs puisqu'ils ne sont pas partie prenante de leur conception.

5.3 L'efficacité

Ce critère repose sur la possibilité de porter un jugement d'évaluation équitable envers le stagiaire à partir des outils prescrits. Il est traité en deux temps, soit du point de vue formatif par l'accompagnement et du point de vue certificatif.

Du point de vue formatif, outre les différents points abordés précédemment concernant spécifiquement les grilles, il a semblé important de comprendre le processus entourant la prise de décision des participants au regard du cheminement des stagiaires dans leur programme de formation initiale. Dans cette optique, les participants se sont prononcés sur **l'importance de connaître les éléments qui font l'objet d'évaluation dans les stages précédents et/ou suivants** ainsi que **les cours suivis à l'université durant toute la formation**. Ils pouvaient consulter deux documents, un tableau qui montre les éléments d'une compétence qui feront l'objet d'évaluation spécifique à chacun des quatre stages et une grille de cheminement des cours suivis lors de chacune des années de la formation initiale.

La grande majorité, tant du primaire que du secondaire, veut connaître les exigences des stages suivants et précédents car cela permet de comprendre la progression des apprentissages du stagiaire, d'être informé sur les exigences que les superviseurs peuvent avoir à leur égard ainsi que d'avoir une idée plus précise de ce qui peut être exigé de lui dans un stage futur.

- *Connaître les éléments qui font l'objet d'évaluation aux stages précédents ou suivants peut me permettre d'avoir des indices concernant ce que mon stagiaire doit atteindre dans l'immédiat sans lui en demander trop (MS36).*

Par contre, en ce qui concerne la grille de cheminement des cours suivis par le stagiaire, si les 2/3 des enseignants du primaire se montrent intéressés, seulement la moitié des enseignants du secondaire sont d'avis qu'il est utile de la connaître. Ces derniers semblent plus centrés sur le niveau d'atteinte des objectifs du stage, tandis que les enseignants du primaire semblent plus à l'aise de porter un jugement lorsqu'ils peuvent mettre le développement des compétences en contexte.

- *On peut faire référence aux cours suivis en cas de difficultés (FP30).*
- *Les noms des cours ne me disent pas le contenu réel et ce n'est pas utile à la réalité du stage. (FS14)*

Ces résultats semblent confirmer ceux de Boutet et Villemin (2014) qui ont remarqué qu'il y a très peu de traces explicites de lien entre cours et stages dans les rencontres en triade et que la modélisation rigoureuse des propositions théoriques est peu présente. Ceci a donc pour effet que le formateur de terrain voit davantage son rôle comme complémentaire à celui du formateur universitaire et ce, en mettant plus l'accent sur le développement de l'identité professionnelle du stagiaire plutôt que sur la reconduction « de ce qui a été appris à l'université ». Ce parallélisme ressenti entre les buts relatifs à l'apprentissage à l'université et ceux relatifs au stage constituerait, selon Feiman-Nemser & Buchmann (1985), le principal piège d'une formation dans laquelle l'articulation théorie-pratique se voudrait intégrative.

Du point de vue certificatif, l'efficacité se traduit par **la compréhension des différentes échelles de mesure et de niveaux de compétences** proposées par les universités. Pour ce faire, trois échelles sont présentées en lien avec chacun des quatre stages de la formation à l'enseignement.

La première qui indique seulement la mention réussite/échec, met en perspective que la décision relative à la réussite du stage est tributaire de l'ensemble des compétences développées par le stagiaire. Il est à noter que dans la plupart des programmes, cette mention est assortie de règles indiquant les exigences relatives à la réussite d'un stagiaire.

La deuxième présente un tableau qui demande au formateur de terrain de donner une note pondérée par compétence, suivi d'une note globale au stage, le tout sur 100%. Cette manière de compiler les notes et de pondérer les compétences questionne le bien-fondé d'une telle approche dans un programme essentiellement porté sur les compétences professionnelles. Stable et recherché du point de vue technique par les personnes habituées à noter les élèves, ce type de document envoie d'abord le message que certaines compétences sont plus importantes que d'autres et qu'ensuite, c'est la somme des 12 compétences, plutôt que leur amalgame, qui permettra de statuer sur la capacité d'enseigner du stagiaire.

La troisième présente un tableau pour chaque compétence dans lequel le formateur de terrain doit indiquer une cote. Le tout est ensuite suivi d'une cote globale pour le stage (allant de A à E). Ce type de grille semble plus en conformité avec une approche par compétences mais elle peut laisser place à interprétation dans la manière de formuler ce qui détermine les cotes.

Bien que la plupart des participants trouvent que les premiers stages peuvent être certifiés par une mention de réussite/échec, les choix se divisent entre ceux du primaire et ceux du secondaire pour la décision finale lors des stages III et IV. De fait, les participants ont d'emblée préféré l'outil qui ressemblait le plus à ce qu'ils complètent habituellement tant auprès de leurs élèves que pour leurs stagiaires. Les participants du primaire rejettent en très grande majorité (85%) la grille avec les notes chiffrées, tandis qu'elle est préférée en totalité par ceux du secondaire.

5.4 Verbatim du primaire :

- *Je préfère mettre réussite et échec pour chaque compétence et la suite appartient à l'université (FP25).*
- *C'est mieux la cote car le but c'est que le stagiaire s'améliore (FP06).*
- *Les chiffres ne sont pas nécessaires, je préfère le langage (FP24).*

5.5 Verbatim du secondaire :

- *Les notes parlent plus (MS31).*
- *Les cotes laissent une certaine latitude dans l'évaluation (FS15)*
- *J'y suis habituée (FS16).*

Finalement et toujours du point de vue certificatif, les participants ont été confrontés à **quatre modèles d'échelles descriptives pour assigner une cote à la compétence CP4**. On remarque que les choix des indicateurs de la compétence sont associés à diverses conceptions théoriques, ce qui peut faire en sorte que le jugement porté par le formateur de terrain soit détourné par le type d'échelle proposé.

La première échelle est élaborée comme un ensemble d'indicateurs de développement allant du dépassement des attentes (niveau 4) à leur non atteinte (niveau 1)⁸. Ce type d'échelles semble répondre à l'idée d'établir des standards pour respecter une forme d'évolution dans les performances attendues du stagiaire (Tardif, 2006). Toutefois cette échelle ne correspond pas réellement à l'établissement de standards au sens de Tardif puisqu'elle est très générale pour l'ensemble des compétences et qu'elle ne définit pas des performances spécifiques à chacun des niveaux pour chacune des compétences.

La deuxième échelle est tout aussi générale. Cinq catégories allant de marqué (a) à très peu développé (e)⁹ sont décrites globalement pour toutes les compétences. Par contre, cette description ne permet pas de juger de la compétence du stagiaire mais vise plutôt la prédiction de sa réussite ou non au prochain stage.

La troisième échelle est élaborée selon le modèle de rubriques (Tardif, 2006) puisqu'elle se sert des composantes des compétences pour déterminer les indicateurs de niveaux allant de 1 à 4¹⁰. Ce modèle exige une définition complète de ce qui est attendu à chaque niveau pour chaque rubrique (donc composante) évaluée. Elle est présentée comme une échelle dont les rubriques sont analytiques permettant ainsi une prise de décision globale ou holistique de la compétence. Ce type de grille risque toutefois d'être exponentiel et de s'avérer lourd pour un formateur de terrain.

Finalement, la quatrième échelle a utilisé le modèle développé par le MÉLS au début des années 2000 pour identifier les niveaux de compétence des élèves au secondaire¹¹. S'apparentant à une échelle de développement, elle n'en a par contre pas les caractéristiques fondamentales car les niveaux ne sont que l'ajout de mots clés permettant d'affirmer un niveau de complexité lié à une compétence.

Au regard de ces quatre cas de figure, les participants préfèrent d'emblée la deuxième échelle, suivie de très près par la première. Ils trouvent les catégories plus nuancées. La troisième échelle suit de près car les participants trouvent la description claire et plus détaillée. Toutefois, cette échelle a obtenu moins d'appui car elle est perçue comme étant plus complexe et trop centrée sur la progression, ce qui ne permet pas de certifier de la compétence. La quatrième échelle est rejetée.

Il leur a été également demandé de se prononcer sur l'efficacité de ces échelles au regard du seuil de réussite. De fait, elles n'établissent pas la réussite en stage de la même manière et parfois, il est difficile de cerner ce seuil qui semble plutôt implicite. C'est la 4^e échelle qui a été choisie par la grande majorité des répondants car le seuil était clairement indiqué de manière visuelle par un trait foncé au niveau D. Les autres échelles donnaient des indications en mots clés.

On note ici que du point de vue certificatif, l'efficacité des échelles analysées par les participants ne correspond pas aux dimensions conceptuelles et théoriques mais plutôt à la logistique organisationnelle et à leur haut niveau technologique. La notation privilégiée est celle qui correspond au type de notation qui prime au primaire ou au secondaire et l'échelle efficace sera claire, visuelle, ciblée et ce, peu importe si elle représente bien ou non des rubriques ou des niveaux de développement. On peut donc supposer qu'encore une fois, le

⁸ Exemple échelle 1 : Manifeste un excellent niveau professionnel (niveau 4)

⁹ Exemple échelle 2 : Prêt à passer au prochain stage (niveau a)

¹⁰ Exemple échelle 3 : Présente aux élèves des moyens et des ressources variées pour la réalisation de leurs tâches, les guide efficacement dans leurs utilisations et suggère d'autres avenues à ceux qui éprouvent des difficultés à les utiliser (niveau 4).

¹¹ Exemple échelle 4 : toute (A), majorité (B), plusieurs (C), peu (D) et peu (E).

défi des concepteurs est d'éviter le piège illustré plus tôt et d'allier à la fois les éléments conceptuels et la facture visuelle de l'échelle pour faciliter un jugement équitable de la part des formateurs de terrain.

6. Conclusion

Cet article avait pour but de donner la parole à des formateurs de terrain afin qu'ils analysent des outils d'accompagnement et d'évaluation au regard d'une prise de décision équitable quant aux niveaux de compétences développés par le stagiaire à travers les critères d'efficacité, de pertinence cognitive et d'efficacité. Des différences sont apparentes entre les enseignants du primaire et ceux du secondaire, notamment sur la manière de noter les compétences et sur l'importance de connaître la grille de cheminement des stagiaires pour porter un jugement équitable. La plupart des participants aiment écrire des commentaires puisqu'ils ont ainsi l'impression d'émettre un point de vue dans un contexte particulier. Tous s'entendent pour affirmer qu'une grille se doit d'être facile d'utilisation mais on remarque que cette facilité ne se traduit pas toujours de la même manière et surtout, qu'elle n'est pas tributaire d'une conception pédagogique ou encore, obligatoirement en lien avec une approche par compétence.

Les résultats de cette recherche évaluative ont ainsi soulevé des réflexions sur l'impact de tels choix d'outils et d'échelles sur l'équité envers le stagiaire et sur les répercussions qu'ils peuvent avoir sur leur avenir ainsi que sur la persévérance dans leurs études (Label, Bélaïr et Goyette, 2012). À cet égard, lors d'une rencontre informelle (rencontre de concertation, mai 2014), un superviseur universitaire ayant à travailler avec des grilles de différents programmes a mentionné qu'il a dû faire échouer un stagiaire à cause de la grille prescrite et que, si ce stagiaire avait été dans un autre programme, il aurait réussi son stage. À ceci s'ajoutent des propos de participants sur l'utilisation de grilles d'autres universités jugées plus faciles, pour évaluer un stagiaire ; ou encore, sur le fait qu'il est difficile de faire échouer un stagiaire et de le lui annoncer, ce qui influence la note donnée ; ou finalement sur l'impossibilité d'inscrire un A+ puisque « personne n'est parfait ».

Au vu de l'analyse de ces verbatim et à l'instar des recommandations de l'OCDE (2002), on peut faire l'hypothèse que les conceptions éducatives sous-jacentes aux outils prescrits et les contextes particuliers d'utilisation en stage par les formateurs de terrain induisent, intentionnellement ou non, des iniquités envers les stagiaires. Dans ce compromis entre réalités et épistémologies, entre prescrit et réel, il s'agit de créer une mouvance dans l'articulation et la mise en œuvre des outils, de telle sorte que ces derniers évoluent au fil des discussions, échanges et formations et ce, pour que l'écart entre le « faisable » et le « idéal » soit vu comme un espace d'apprentissage et de développement (Mayen, 2008, Tessaro, 2015). Tout ceci amène à penser qu'un travail de fond misant entre autres sur les communautés de pratiques (Label, Bélaïr et Ducharme, 2015) doit se poursuivre au niveau de l'impact des outils prescrits dans les différents programmes.

7. Références

- Allal, L. & Mottier Lopez, L. (2007). *Régulation des apprentissages en situation scolaire et en formation*. Bruxelles : De Boeck.
- Bates, A. J. & Burbank, M. D. (2008). Effective Student Teacher Supervision in the Era of "No Child Left Behind". *Professional Educator*, 32(2), 1-11.
- Bélaïr, L. M. (2005). Les outils d'observation. In N. Rousseau (Ed.), *Se former pour mieux superviser* (pp. 125-144). Montréal : Guérin.

- Bélaïr, L. M. (2007). Défis et obstacles dans l'évaluation des compétences professionnelles. In L. M. Bélaïr, D. Laveault & C. Lebel (Eds.), *Les compétences professionnelles en enseignement et leur évaluation* (pp. 181-196). Ottawa : PUO.
- Bélaïr, L.M. & Baribeau, A. (2010). Analyse des dispositifs de régulation dans un cours en évaluation des apprentissages. In L.M. Bélaïr, C. Lebel, N. Sorin, A. Roy & L. Lafortune (Dir.). *Régulation et évaluation des compétences en enseignement : vers la professionnalisation* (pp. 223-236). Montréal : PUQ.
- Bélaïr, L.M. & Dionne, É. (2010). Évaluation des apprentissages dans le contexte québécois : entre décisions politiques et pratiques en salle de classe. *Mesure et Évaluation en Éducation*, 32(3), 47-76.
- Bélaïr, L.M. (2014). Évaluer ce qu'ils ont appris. In L. Ménard, & L. St-Pierre (Dir.). *Se former à la pédagogie de l'enseignement supérieur* (pp. 355-380). Montréal : Chenelière Éducation.
- Bélaïr, L.M., Vivegnis, I. & Lafrance, J. (2015). Évaluation de la mise en œuvre d'un dispositif de formation d'enseignants associés au Québec par le biais d'une recherche développement. *Revue RİPES*, 31(1), <https://ripes.revues.org/925>.
- Belna, K., Leménager, T. & Mermet, L. (2012). Evaluer l'efficacité de politiques et programmes environnementaux : quel cadrage méthodologique proposer ? Le cas du Fonds de partenariat pour le carbone forestier. *Notes méthodologiques*, 5. Paris : Division Évaluation et capitalisation, Agence Française de Développement.
- Boutet, M. & Villemin, R. (2014). L'accompagnement : un élément clé pour l'apprentissage en stage et pour le développement professionnel continu des enseignants, *Éducation et socialisation* [En ligne], 35, 1-12.
- Boyer, L. (2013). *Coconstruction d'un modèle cognitif de l'apprentissage d'une compétence en vue d'assurer la validité et l'équité de son évaluation : Le cas de la compétence « Exercer un jugement clinique infirmier »* (Thèse de doctorat). Université de Sherbrooke, Québec.
- Brochu, A.-C. & Bouvier, F. (2008). Évaluer en même temps que les compétences se développent. In L. Lafortune, S. Ouellet, C. Lebel & D. Martin (Dir.), *Réfléchir pour évaluer des compétences professionnelles à l'enseignement : deux regards, l'un québécois et l'autre suisse* (pp. 109-133). Québec : Presses de l'Université du Québec.
- Camaraire, L. & Cournoyer, É. (2002). L'évaluation des prestations des stagiaires – les perspectives des évaluateurs et des évalués, convergences ou divergences ? In G. Boutin (Dir.), *Formation pratique des enseignants et partenariat : état des lieux et prospective* (pp. 162-177). Montréal : Éditions Nouvelles.
- De Ketele, J.-M. (2000). En guise de synthèse : Convergences autour des compétences. In C. Bosman, F. M. Gerard, & X. Roegiers (Dir), *Quel avenir pour les compétences* (pp. 187-192). Bruxelles : De Boeck Université.
- Desbiens, J.-F., Spallanzani, C. & Borges, C. (2013). *Quand le stage en enseignement déraile : regards pluriels sur une réalité trop souvent occultée*. Québec : Presses de l'Université du Québec.
- Feiman-Nemser, S. & Buchmann, M. (1985). Pitfalls of experience in teacher preparation. *The Teachers College Record*, 87(1), 53-65.
- Figari, G. (2013). L'évaluation entre « technicité » et « théorisation » ? *Mesure et Évaluation en Éducation*, 36(3), 5-24.
- Figari, G. & Remaud, D. (2014). *Méthodologie d'évaluation en éducation et formation, ou l'enquête évaluative*. Bruxelles : De Boeck.
- Gagnon, C. & Mazalon, É. (2008). *L'approche qualité dans la formation à l'enseignement professionnel en alternance : la nécessaire collaboration des acteurs*. Actes du colloque du 25e congrès de l'Association internationale de pédagogie universitaire, Montpellier, 2008.
- Gauthier, C. & Mellouki, M. (2006). *La formation des enseignants au Québec à la croisée des chemins : conditions et promesses de l'approche de formation par compétences*. Sainte-Foy : Les Presses de l'Université Laval.
- Gervais, C. & Lepage, M. (2000). Transfert de la responsabilité de l'évaluation en stage du superviseur vers l'enseignant associé : un pas de plus vers la professionnalisation. In D. Martin, C. Gervais, C. Garant & C. St-Jarre, (Dir.), *Recherches et pratiques en formation des maîtres. Vers une pratique réfléchie et argumentée* (pp. 113-128). Sherbrooke : CRP.

- Gouin, J.-A. & Hamel, C. (2015). La perception de formateurs de stagiaires quant au développement et à l'évaluation formative des quatre compétences liées à l'acte d'enseigner. *Revue Canadienne d'Éducation*, 38(3), 27p. <http://journals.sfu.ca/cje/index.php/cje-rce/issue/view/106>
- Knight, P. (2006). The Local Practices of Assessment. *Assessment & Evaluation in Higher Education*, 31(4), 435-452.
- Lapointe, J.-R. & Guillemette, F. (2015). L'évaluation des stages par les acteurs de la formation pratique : modalités, supervision, évaluation et guide de stage. *Revue des sciences de l'éducation*, 41(2), 199-217.
- Label, C., Bélaïr, L.M. & Goyette, N. (2012). Accompagnement et reconnaissance professionnelle au service de la persévérance des stagiaires en difficulté. *Revue Recherches et Éducation*. 7, 55-68
- Label, C., Bélaïr, L.M. & Ducharme, N. (2015). *Analyse des postures dans l'accompagnement et l'évaluation par des formateurs de terrain*, Communication à l'AFORME, Québec.
- Label, C., Bélaïr, L.M. & Monfette, O. (2015). Pratiques de jugement d'évaluation d'enseignants associés auprès de leurs stagiaires. In P-F. Coen & L.M. Bélaïr, *Évaluation et autoévaluation, quels espaces de formation ?* (pp. 157-172). Bruxelles : De Boeck.
- Label, C., Bélaïr, L.M., Monfette, O., Hurtel, B., Miron, G. & Blanchette, S. (2016). Formation de stagiaires en enseignement en situation de handicap : points de vue des formateurs de terrain. *Revue Éducation et francophonie*, 44(1).
- Leroux, J.-L. & Bélaïr, L.M. (2015). Exercer son jugement professionnel en enseignement supérieur. In J.-L. Leroux (Dir.), *Évaluer les compétences au collégial et à l'université : un guide pratique* (pp. 67-107). Sherbrooke : Association québécoise de pédagogie collégiale (coll. PERFORMA).
- Leroux, J.-L. (2014). Évaluer pour faire apprendre. In L. Ménard, & L. St-Pierre (Dir.). *Se former à la pédagogie de l'enseignement supérieur* (pp. 321-354). Montréal : Chenelière Éducation.
- Leshem, S. & Bar-Hama, R. (2008). Evaluating Teaching Practice. *ELT Journal*, 62(3), 257-265.
- Mayen, P. (2008). Intégrer les savoirs à l'action. In P. Perrenoud, M. Altet, C. Lessard & L. Paquay (Éds), *Conflits de savoirs en formation des enseignants* (pp. 43-58). Bruxelles : De Boeck.
- Miron, G. & Blanchette, S. (2015). *Analyse des grilles de stage par des enseignants associés*. Communication présentée dans le cadre du séminaire REÉVES-GRAPPE, Bruxelles.
- Mizzaro, S. (1997). Relevance: the whole history. *Journal of the American society for information science*, 48(9), 810-832.
- Monnard, I. & Luisoni, M. (2013). Sens attribué à l'évaluation des compétences professionnelles par tâches complexes chez de futurs enseignants en formation. *Mesure et évaluation en éducation*, 36(2), 1-27.
- Nguyen, D.-Q. & Blais, J.-G. (2007). Approche par objectifs ou approche par compétences ? repères conceptuels et implications pour les activités d'enseignement, d'apprentissage et d'évaluation au cours de la formation clinique. *Pédagogie médicale*, <http://www.pedagogie-medicale.org>
- OCDE (2002). *Glossary of Key Terms in Evaluation and Results Based Management*. Paris: OECD.
- Ouellet, S., Pache, A. & Portelance, L. (2008). L'encadrement, l'évaluation et la concertation au service de la formation des stagiaires en enseignement. In L. Lafortune, S. Ouellet, C. Label & D. Martin (Dir.), *Réfléchir pour évaluer des compétences professionnelles à l'enseignement : deux regards, l'un québécois et l'autre suisse* (pp. 59-78). Québec : Presses de l'Université du Québec.
- Perrenoud, Ph. (2001). Dix défis pour les formateurs d'enseignants. In Ph. Perrenoud, *Développer la pratique réflexive dans le métier d'enseignant*. Professionnalisation et raison pédagogique. Paris : ESF Éditeur.
- Portelance, L. & Tremblay, F. (2006). Les responsabilités complémentaires de l'enseignant associé et du stagiaire au regard de la formation à l'enseignement. In J. Loisel, L. Lafortune & N. Rousseau (Dir.), *L'innovation en formation à l'enseignement. Pistes de réflexion et d'action* (pp. 41-54). Québec : Presses de l'Université du Québec.
- Portelance, L. (2010). Analyse des perceptions du soutien d'un enseignant associé à la formation du stagiaire, *Éducation et francophonie*, 38(2), 21-38.
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Bruxelles : De Boeck.
- Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation, Volume I* (pp. 39-83). Chicago, Il.: Rand McNally.

- Simonnot, B. (2008). La pertinence en sciences de l'information : des modèles, une théorie ? In F. Papy (Dir.), *Problématiques émergentes dans les Sciences de l'Information* (pp. 161-182). Paris : Hermes-Lavoisier.
- Stufflebeam, D. (2001). Evaluation models. *New directions for evaluation*, (89), 7-98.
- Tang, S.Y. (2008). Issues in the Field Experience Assessment in Teacher Education in a Standards Based Context. *Journal of education for Teaching*, 34(1), 17-32.
- Tardif, J. (2006). *L'évaluation des compétences*. Montréal : Chenelière Éducation.
- Tessaro, W. (2015). Formation permanente des enseignants à l'évaluation des apprentissages : une continuité à construire. In L. Bélair & P.-F. Coen (Eds.), *Évaluation et autoévaluation. Quels espaces de formation ?* (pp. 55-69). Bruxelles : De Boeck.
- Tricot, A. & Tricot, M. (2000). Un cadre formel pour interpréter les liens entre utilisabilité et utilité des systèmes d'information. *Actes du 12^e colloque ERGO-IHM*, Biarritz, 195-202. Récupéré le 17 avril 2010 de http://andre.tricot.pagesperso-orange.fr/TricotTricot_Ergo_IHM.pdf
- Van Nieuwenhoven, C. & Colognesi, S. (2013). Une recherche collaborative autour des difficultés des maîtres de stage à accompagner leur stagiaire. *Interações*, 9(27), 118-138.
- Vedung, E. (2008). *Public Policy and Program Evaluation*. Nouveau-Brunswick: Transaction Publishers.