

## Sens attribué à l'évaluation des compétences professionnelles par tâches complexes chez de futurs enseignants en formation

Isabelle Monnard et Marc Luisoni

Volume 36, numéro 2, 2013

Date de réception : 5 novembre 2012

Date de réception de la version finale : 14 août 2013

Date d'acceptation : 10 septembre 2013

URI : <https://id.erudit.org/iderudit/1024413ar>

DOI : <https://doi.org/10.7202/1024413ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (imprimé)

2368-2000 (numérique)

[Découvrir la revue](#)

Citer cet article

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.  
<https://doi.org/10.7202/1024413ar>

Résumé de l'article

Depuis plusieurs années, la formation des enseignants en Suisse a vu la mise en oeuvre de curricula basés sur une logique d'acquisition de compétences professionnelles. La Haute École pédagogique de Fribourg certifie le parcours de ses étudiants par le biais d'évaluations par tâches complexes réparties au cours de la formation de niveau Baccalauréat. La recherche présentée s'intéresse aux représentations développées par les étudiants à propos de ce mode d'évaluation, et plus précisément au sens qu'ils attribuent à cette modalité d'évaluation rarement rencontrée dans leur parcours scolaire. L'analyse qualitative met en évidence un mode d'évaluation perçu comme fortement producteur de sens malgré une connaissance souvent lacunaire des critères d'évaluation institutionnels. La mise en évidence d'apprentissages réalisés par les étudiants dans ce cadre suscite un questionnement quant à la validité des tâches proposées, ainsi qu'une réflexion au sujet de la frontière entre évaluation formative et évaluation sommative, dans un contexte d'évaluation par tâches complexes.

## Sens attribué à l'évaluation des compétences professionnelles par tâches complexes chez de futurs enseignants en formation

Isabelle Monnard

Marc Luisoni

Haute École pédagogique de Fribourg

MOTS CLÉS: évaluation, compétence, tâche complexe, formation des enseignants, sens

*Depuis plusieurs années, la formation des enseignants en Suisse a vu la mise en œuvre de curricula basés sur une logique d'acquisition de compétences professionnelles. La Haute École pédagogique de Fribourg certifie le parcours de ses étudiants par le biais d'évaluations par tâches complexes réparties au cours de la formation de niveau Baccalauréat. La recherche présentée s'intéresse aux représentations développées par les étudiants à propos de ce mode d'évaluation, et plus précisément au sens qu'ils attribuent à cette modalité d'évaluation rarement rencontrée dans leur parcours scolaire. L'analyse qualitative met en évidence un mode d'évaluation perçu comme fortement producteur de sens malgré une connaissance souvent lacunaire des critères d'évaluation institutionnels. La mise en évidence d'apprentissages réalisés par les étudiants dans ce cadre suscite un questionnement quant à la validité des tâches proposées, ainsi qu'une réflexion au sujet de la frontière entre évaluation formative et évaluation sommative, dans un contexte d'évaluation par tâches complexes.*

KEY WORDS: evaluation, assessment, complex tasks, professionals skills, teachers education, meaning

*For several years, curricula based on the acquisition of professional skills were implemented in the training of teachers in Switzerland. The University of Teachers Education of Fribourg certifies students through assessments with complex tasks throughout the three years of training. Our research focuses on representations of students about this form of assessment, and more specifically on the meaning they attribute to this innovative process, which they rarely encountered in their*

*previous school careers. The qualitative analysis reveals that this assessment method is perceived as meaningful, despite the fact that students often lack knowledge of institutional assessment criteria. Furthermore, it suggests that learning achieved by students in this context raises questions as to the validity of the proposed tasks. Finally, we discuss its implications on the distinction between formative and summative assessment in the context of assessment of complex tasks.*

**PALAVRAS-CHAVE:** avaliação, competências, tarefa complexa, formação de professores, sentido

*Durante muitos anos, a formação de professores na Suíça assistiu à implementação de currículos baseados numa lógica de aquisição de competências profissionais. A Escola Superior de Educação de Friburgo certifica o percurso destes estudantes através da avaliação das tarefas complexas distribuídas ao longo do Bacharelato. A investigação aqui apresentada debruça-se sobre as representações construídas pelos estudantes sobre este modo de avaliação e, mais precisamente, sobre o sentido que eles atribuem a esta modalidade de avaliação pouco comum no seu percurso escolar. A análise qualitativa coloca em evidência um modo de avaliação que é percebido como sendo fortemente produtor de sentido, apesar da ausência de um conhecimento dos critérios de avaliação institucionais. A evidenciação das aprendizagens realizadas pelos estudantes neste quadro suscita um questionamento sobre a validade das tarefas propostas, bem como uma reflexão sobre a fronteira entre a avaliação formativa e a avaliação sumativa num contexto de avaliação por tarefas complexas.*

---

Note des auteurs – Pour faciliter la lecture du texte, le masculin générique est utilisé pour désigner les deux sexes. La correspondance liée à cet article peut être adressée à Isabelle Monnard, Haute École pédagogique de Fribourg, rue de Morat 36, CH – 1700 Fribourg, téléphone +41 26 305 71 39, ou à l'adresse courriel suivante : [monnardi@eduf.ch]

## **Introduction**

Depuis une dizaine d'années, avec la création des hautes écoles pédagogiques en Suisse, la formation des enseignants a été profondément modifiée afin de répondre aux exigences d'une formation de niveau tertiaire. La logique curriculaire classique qui prévalait au sein des écoles normales a été abandonnée au profit d'une approche par compétences professionnelles développées au sein d'un référentiel.

Ce changement de dispositifs de formation des enseignants a eu pour effet, notamment, d'obliger les institutions et les formateurs à repenser en profondeur les modalités d'évaluation. Dans une logique de compétences, il ne s'agit en effet plus d'évaluer des savoirs et des savoir-faire disparates et décontextualisés, mais bien leur mobilisation pour résoudre des tâches significatives.

La Haute École Pédagogique de Fribourg (HEP FR) a opté dès son ouverture pour une évaluation certificative des compétences au moyen de tâches complexes. Ce changement de paradigme évaluatif a été accompagné d'un dispositif de suivi afin d'en garantir la qualité. Différents travaux ont été réalisés dans ce cadre (Coen, Galland, Monnard, Perrin, & Rouiller, 2008 ; Galland, Monnard, & Coen, 2010). Cette recherche s'inscrit dans cette perspective en s'intéressant aux perceptions de ce mode d'évaluation par les étudiants.

Après une brève présentation des concepts ayant servi à développer le dispositif d'évaluation au sein de la HEP FR ainsi que le contexte de sa mise en œuvre, les résultats issus d'analyses d'entretiens réalisés auprès des étudiants sont présentés.

## **Les compétences et leur évaluation**

La question de la compétence est largement traitée dans la littérature et les définitions sont nombreuses (Allal, 2002 ; Aubret & Gilbert, 2003 ; Perrenoud, 1999 ; Tardif, 2006), d'aucuns allant même jusqu'à la considérer comme une énigme (Dolz & Ollagnier, 2002). Dans le cadre de cette étude, les auteurs se réfèrent à la définition largement partagée de Roegiers et De Ketele (2000) pour qui la compétence est « la possibilité, pour un individu, de mobiliser de manière intériorisée un ensemble intégré de ressources en vue de résoudre

une famille de situation problème» (p. 66). Comme le suggère Le Boterf (2006), il ne suffit donc pas qu'un professionnel possède les ressources pour être considéré comme compétent, mais bien qu'il soit capable de les mobiliser de manière pertinente en situation professionnelle.

Appuyer une formation professionnelle des enseignants sur un référentiel de compétences implique donc bien, au-delà de l'acquisition de savoirs disciplinaires et didactiques et de gestes professionnels, un travail sur des situations nécessitant la mobilisation et l'articulation de ces ressources. Ces situations doivent permettre de développer les compétences et d'inférer en fin de période ou de formation le niveau d'atteinte minimum, si tant est que la compétence soit définie dans une perspective de continuum (Allal, 2002). Dans cette logique, l'évaluation se trouve naturellement profondément modifiée puisqu'il s'agit de s'assurer non seulement de l'acquisition des ressources, mais bien de vérifier que les étudiants ont acquis les compétences leur permettant de résoudre des situations authentiques en mobilisant un ensemble de ressources à leur disposition. Les formateurs sont donc amenés à développer ces situations censées permettre l'évaluation des compétences. Scallon (2004) met en évidence le double enjeu de tout formateur confronté à la création d'un tel dispositif : premièrement, la conception de situations dans lesquelles les apprenants seront placés afin qu'ils puissent démontrer ce dont ils sont capables, et, deuxièmement, la construction d'outils de jugement « permettant d'intégrer en un tout cohérent divers indices observés au cours ou au terme de la progression des élèves » (p. 102).

Utiliser ce mode d'évaluation pour répondre à ces exigences implique, comme pour tout type d'évaluation, d'en vérifier les qualités. Si l'approche psychométrique classique ne peut être utilisée (Scallon, 2004), s'assurer que les tâches proposées respectent certains principes fondamentaux permet néanmoins de garantir la fiabilité du dispositif (Beckers, 2002 ; Roegiers, 1999, 2003).

Plusieurs caractéristiques participent à la validité des tâches complexes. En premier lieu, il s'agit de s'assurer du sens des situations par rapport à la suite de la formation ou à la vie professionnelle, autrement dit leur caractère authentique ou réel ainsi que leur appartenance au curriculum de formation. Il faut ensuite vérifier si les situations sollicitent effectivement des compétences et mesurent ce qu'elles prétendent mesurer. Afin de ne pas rester sur le plan de la restitution ou de la répétition, les situations doivent être inédites et permettre des solutions originales. Toutefois, les auteurs ne suivent pas Crahay

(2006) sur ce qui constitue le caractère inédit d'une situation. En effet, de par leur nature même, les situations éducatives sont souvent uniques et donc inédites ; il n'est donc pas nécessaire qu'elles présentent un caractère extrême ou exceptionnel. Et si, comme le précise Gérard (2009), être compétent, «c'est savoir gérer l'inédit lorsqu'il survient» (p. 11), l'important réside moins dans l'aspect inédit de la situation que dans l'originalité des solutions possibles, originalité devant être entendue comme l'adaptation au caractère unique de la situation éducative. Sur le plan de la construction, il est important également que les consignes ne donnent pas ou peu d'indications sur les ressources à mobiliser (Beckers, 2002 ; Scallon, 2004). La compétence consistant justement dans la mobilisation et la combinaison des ressources adéquates, il revient à l'étudiant de choisir celles qui seront pertinentes pour faire face à une situation. En ce qui concerne les contenus, les situations doivent être en cohérence avec ce qui a été enseigné, tant sur le plan des ressources que de leur mobilisation dans des situations quasi isomorphes (Jonnaert, 2011).

Lorsque de telles tâches sont utilisées en vue de certifier les apprentissages, les formateurs doivent produire un jugement professionnel en s'appuyant sur des référentiels multiples (Mottier Lopez, 2009) : contraintes institutionnelles, valeurs personnelles, représentations de l'évaluation, culture de l'établissement, interactions avec les collègues, nature de la tâche. Pour permettre un jugement fiable au vu d'une telle complexité, le recours à une évaluation critériée est essentiel (Beckers, 2002). À la HEP FR, chaque tâche est corrigée par des correcteurs différents qui s'appuient sur une évaluation critériée permettant de réduire les écarts intercorrecteurs. Un nombre restreint de critères par tâche (4 à 5) devrait rendre cette évaluation plus «juste» (Roegiers, 2004). En effet, la prise en compte d'un nombre important de critères entraîne un fort risque d'interdépendance entre eux ainsi qu'une surcharge cognitive pour les correcteurs et les étudiants.

Pour garantir la fiabilité du dispositif d'évaluation, il est indispensable de vérifier le respect des principes énoncés ci-dessus. Dans cette optique, le recours à des procédures de validation variées impliquant des recueils de données *a priori* et *a posteriori* auprès des différents acteurs concernés est nécessaire. Les auteurs suivent De Ketele et Gérard (2005) lorsqu'ils mettent en évidence l'importance, «à l'image de l'enquêteur sans preuve flagrante, d'adopter une série de stratégies qui constituent un faisceau d'indices suffisants de la valeur de nos épreuves d'évaluation (à la fois pertinentes, valides et fiables)» (p. 22).

## Contextualisation

À la HEP FR, les différentes étapes de la formation sont certifiées au moyen de sept tâches (ou situations) complexes réparties sur les trois années de formation (voir tableau 1). Les tâches de troisième année font partie des épreuves de la certification finale.

Tableau 1

### *Descriptif des tâches complexes certificatives à la HEP FR*

Année de formation	Code TC	Descriptif
1 <sup>re</sup> année	TC1A	<i>Analyse d'un processus d'apprentissage</i> À partir d'une séquence vidéo d'enseignement-apprentissage, des duos d'étudiants analysent un processus d'apprentissage et mettent en évidence les enjeux cognitifs et sociaux.
	TC1B	<i>Analyse d'une situation de classe</i> Individuellement, sur la base de la description écrite d'une situation de classe rencontrée par un enseignant, les étudiants produisent une analyse à partir des points de vue de la sociologie, de la psychologie et de la pédagogie.
	TC1C	<i>Réflexion sur la profession et les rôles de l'enseignant ou de l'enseignante</i> Après avoir choisi un thème autour des représentations du métier et des rôles de l'enseignant dans le système scolaire, les étudiants, par groupe de quatre, produisent un film vidéo ainsi qu'une synthèse théorique.
2 <sup>e</sup> année	TC2D	<i>Réflexion autour d'un concept pédagogique</i> Par deux, les étudiants mettent en évidence la perspective historique du développement d'un concept pédagogique et confrontent leurs représentations à son propos.
	TC2E	<i>Planification et conduite de séquences d'enseignement-apprentissage</i> Chaque étudiant planifie, et conduit en stage des séquences d'enseignement-apprentissage. Il les analyse <i>a posteriori</i> soit en fonction des concepts d'une didactique spécifique, soit en référence à un concept didactique (mise en commun, consignes, phases, etc.).
3 <sup>e</sup> année	TC3F	<i>Analyse d'une situation de classe</i> À partir de la description d'une situation problématique en lien avec sa profession, l'étudiant élabore une analyse et un plan d'intervention en se référant à ses savoirs théoriques et d'expérience.

TC3G *Planification conduite et analyse de séquences d'enseignement-apprentissage*

Dans une discipline donnée, l'étudiant planifie et élabore un projet d'enseignement-apprentissage qui s'inscrit dans la durée. Il conduit son projet durant le stage de diplôme puis l'analyse sous différents aspects avant de le soutenir devant un jury.

---

Ces tâches doivent permettre de «déterminer le niveau d'acquisition d'une compétence (d'intégration) à travers la mobilisation de savoirs ou de savoir-faire déterminés et en se référant à des critères adéquats de manière à ce que les résultats entraînent la prise de décision adéquate» (Gérard, 2009, p. 59). Les réflexions autour de la culture de l'évaluation dans l'institution ont amené les formateurs à choisir un nombre limité de critères indépendants et identiques pour toutes les tâches. Sur la base des propositions de Roegiers (2004) et de Beckers (2002), quatre critères ont été retenus. Le premier concerne la pertinence des idées, c'est-à-dire l'adéquation du travail avec la tâche demandée et l'intérêt de la production. Le second critère vise à évaluer la cohérence à travers l'organisation des idées et la logique interne. Le troisième critère est ciblé sur la qualité des éléments présentés, la justesse des concepts en regard des disciplines concernées. À noter que, pour la tâche E de 2<sup>e</sup> année, le critère qualité a été scindé en deux, les formateurs souhaitant évaluer séparément la qualité du projet développé et celle de l'analyse réalisée. Finalement, le respect des normes de la langue et de la tâche est évalué dans le critère forme. Les critères choisis, de par leur généralité, ne permettent pas à eux seuls de porter un jugement précis sur les différentes tâches soumises aux étudiants durant leur formation. Ils ont donc été opérationnalisés à l'aide d'indicateurs plus ou moins détaillés et contextualisés en fonction des tâches. Le tableau 2 montre un exemple d'opérationnalisation des quatre critères d'évaluation pour l'une des tâches complexes (TC1B – deuxième tâche de première année : analyse d'une situation de classe d'un point de vue pédagogique, psychologique et sociologique).



Tableau 2  
*Critères et indicateurs de la tâche complexe 1B*

Critères	Indicateurs généraux et particuliers (pour le critère qualité)
<b>Pertinence</b>	Thème réellement traité Travail intéressant, convaincant, original et créatif Questions d'analyse porteuses de sens et pertinentes
<b>Cohérence</b>	Logique interne, liens entre les parties du travail Enchaînements qui font sens Agencement des arguments
<b>Qualité</b>	Description objective <ul style="list-style-type: none"> <li>• distinction entre description et interprétation</li> </ul> Savoirs théoriques issus des trois cours <ul style="list-style-type: none"> <li>• travail prenant en compte les trois perspectives</li> <li>• choix de références pertinentes</li> </ul> Termes/concepts spécifiques <ul style="list-style-type: none"> <li>• utilisation de termes spécifiques adéquats</li> <li>• définition précise</li> <li>• articulation entre les termes</li> <li>• liens avec la situation</li> </ul> Plan d'intervention <ul style="list-style-type: none"> <li>• prise en compte des éléments essentiels de la situation</li> <li>• cohérence entre buts et moyens</li> <li>• mise en place de pistes d'action différenciées</li> </ul> Validité des interventions proposées <ul style="list-style-type: none"> <li>• interventions à haute probabilité de réussite</li> </ul>
<b>Forme</b>	Respect du code éthique Respect de la consigne de la tâche Respect des normes de la langue orale et écrite <ul style="list-style-type: none"> <li>• (style, orthographe, syntaxe, ponctuation)</li> </ul> Respect des normes APA

Pour réaliser chacune des tâches, les étudiants disposent de la consigne, des critères et des indicateurs. Cette manière de procéder leur permet de savoir ce qu'on attend d'eux et de mieux concentrer leur attention dans une perspective d'évaluation instituée (Barbier, 1985). La tâche est considérée comme acquise si l'étudiant atteint un seuil minimal sur chacun des critères (quatre niveaux : -, +, ++ et +++). Les indicateurs donnent de l'information pour décider du niveau d'acquisition du critère, mais ils ne doivent pas être évalués un par un.

Les tâches proposées concernent plusieurs dispositifs de formation (cours, ateliers<sup>1</sup>, stage) et des formateurs issus de domaines variés. Dès la mise en place de cette nouvelle forme d'évaluation, des questions ont émergé notamment à propos du lien entre ces tâches et les cours dispensés, de la nature des compétences évaluées, de l'objectivité et de la fiabilité des évaluations réalisées par les formateurs. Les tâches ont donc fait l'objet d'un suivi particulier pour s'assurer de leur qualité.

Une première recherche (Coen et al., 2008) a consisté à analyser *a priori* les descriptifs des tâches afin de vérifier le respect de certains des principes définissant une situation complexe (présence de critères, consigne ouverte, mobilisation de compétences, etc.). Toutefois, proposer des tâches respectant ces principes ne suffit pas à garantir qu'elles évaluent bien ce qu'on souhaite évaluer. Afin de vérifier la pertinence des tâches proposées, les auteurs ont choisi de soumettre leurs descriptifs aux enseignants de terrain en leur demandant de s'exprimer sur leur caractère authentique et les compétences qui, d'après eux, étaient nécessaires pour les effectuer. Ces deux perspectives ont permis de mettre en évidence une bonne adéquation entre les tâches et des pratiques réelles du métier ou de la formation, ainsi que le respect d'un bon nombre de caractéristiques permettant de définir une tâche complexe. Dans un deuxième temps, les auteurs ont soumis les tâches cette fois au regard des formateurs qui les ont conçues, en les interrogeant sur la manière dont ils les ont construites ainsi que sur la perception de leur évolution (Galland et al., 2010). Les résultats ont mis en évidence le caractère non stabilisé de telles tâches et l'importance de les remettre au centre des échanges sur l'évaluation. Cela a permis de montrer la difficulté de modifier durablement les pratiques évaluatives et la nécessité de construire et de maintenir une culture commune autour de l'évaluation par tâches complexes. À travers ces recherches, les auteurs ont également pu mettre en évidence le rôle important des critères dont la compréhension suscite encore des interrogations.

Après avoir questionné, dans d'autres travaux, la pertinence et la validité des tâches au travers du regard des formateurs et des enseignants primaires, les auteurs se focaliseront sur l'évalué. Peu de travaux ont pris en compte la perception de la démarche d'évaluation par les évalués eux-mêmes (Scallon, 2004). Ainsi, si la signification et le réalisme d'une tâche concourent fortement à la qualité, il ne suffit pas que les concepteurs s'accordent sur la validité de contenu pour que la tâche soit perçue comme authentique par les étudiants.

Un autre critère permettant d'accroître la validité consiste à rendre les éléments de la démarche (consigne, critères, etc.) aussi explicites que possible pour les évalués (Rey, 2008). Les interroger sur leur compréhension de la tâche permettra de vérifier le caractère institué du dispositif par la connaissance qu'ont les évalués du processus d'évaluation (Hadji, 2012). Les premières questions de recherche sont les suivantes : quelle perception les futurs enseignants en formation initiale ont-ils du dispositif d'évaluation par tâche complexe ? Quelle compréhension ont-ils des critères ? L'indépendance ou la dépendance des critères identifiée par les étudiants correspond-elle à une réalité ? Quelle peut être l'utilité du retour des formateurs ?

De ce questionnement initial, il a finalement émergé une problématique centrale qui peut se définir en ces termes : Quel sens les étudiants construisent-ils à partir des modalités d'évaluation par tâches complexes ?

## Méthode

La forte implication des auteurs dans le terrain de cette recherche, puisque formateurs au sein de l'institution, les amène privilégier une posture issue de la théorie ancrée (Strauss & Corbin, 2004). Cette étude s'appuie sur une méthodologie essentiellement qualitative et s'inscrit dans une visée exploratoire. Les données recueillies ont été récoltées dans le cadre d'entretiens de type compréhensif retranscrits puis analysés en fonction de catégories émergentes. La posture méthodologique des auteurs invite à concevoir l'analyse des données comme un processus dynamique basé sur un constant aller-retour entre l'élaboration conceptuelle et le terrain (Miles & Huberman, 2003), ce qui implique une évolution du protocole de questionnement en cours d'analyse. Seize entretiens ont été réalisés auprès de sept étudiants de deuxième et neuf de troisième année en formation initiale d'enseignants de degrés préscolaires et primaires. Parmi eux, sept n'avaient pas réussi au moins une tâche lors d'une première tentative.

Lors des entretiens, les étudiants ont, dans un premier temps, été entendus sur le caractère explicite du dispositif d'évaluation. Ils ont été invités à décrire les caractéristiques d'une tâche complexe et de son évaluation, comme ils le feraient par rapport à un étudiant nouvellement arrivé. Il leur a ensuite été demandé de revenir sur une tâche complexe réussie ou échouée à la fin de l'année précédente (TC1B ou TC2E) et de s'exprimer sur la nature de la tâche, sa réalisation ainsi que sur leur ressenti face à l'évaluation et au retour des

formateurs. Pour les sept étudiants de deuxième année, il s'agissait d'évoquer la tâche complexe de première année qui consiste en une analyse de situation proposée par les formateurs à partir des champs de la psychologie, de la pédagogie et de la sociologie de l'éducation (TC1B). Pour les étudiants de troisième année, l'entretien se basait sur une tâche de deuxième année demandant la planification et l'analyse de plusieurs séquences d'enseignement-apprentissage menées en stage (TC2E). Ces deux tâches sont à réaliser individuellement, mais sans contrainte de temps. Pour évaluer ces tâches, les formateurs disposent des quatre critères présentés plus haut (voir tableau 2), accompagnés d'indicateurs plus ou moins détaillés. Pour chaque critère, ils doivent donner une appréciation qui va de «-» à «+++». Si un seul des critères obtient un «-», la tâche est considérée comme non acquise.

Dans un deuxième temps, à partir des TC1B et TC2E et des autres TC rencontrées par les étudiants dans leur parcours à la HEP FR, l'entretien a permis d'aborder plusieurs thématiques : la compréhension des différents critères, la perception du retour de l'évaluation réalisée par les formateurs et les mécanismes d'autorégulation mis en œuvre pour répondre aux attentes de la tâche. Enfin, en vue de questionner la signification et le réalisme du dispositif au regard de la formation et de la pratique future, la pertinence et le sens de cette forme d'évaluation en général ont été investigués.

En complément à cette analyse qualitative et afin de mettre en lien la perception des étudiants et l'évaluation effective, une analyse quantitative des appréciations données par les formateurs à chacun des quatre critères pour les deux tâches de tous les étudiants a été effectuée afin de mesurer l'indépendance des critères qui, selon Roegiers (2004), participe à la validité de l'évaluation.

## Résultats

Dans un premier temps, les propos des étudiants ont été comparés en fonction de la réussite ( $n = 9$ ) ou de l'échec ( $n = 7$ ) à la TC servant de point de départ à l'entretien. Les analyses de contenu effectuées montrant de grandes similitudes entre les deux groupes, les auteurs n'ont pas retenu ce critère dans la présentation des résultats

### ***Perception du dispositif de l'évaluation par tâche complexe***

Afin de déterminer dans quelle mesure les étudiants se sont appropriés le mode d'évaluation adopté par la HEP FR, il leur a été demandé, d'une part, d'identifier les caractéristiques d'une tâche complexe et, d'autre part, de décrire et d'explicitier la manière dont elles sont évaluées par les formateurs.

#### ***Caractéristiques de la tâche complexe***

Quand il est demandé aux étudiants les caractéristiques d'une tâche complexe, leurs propos divergent selon leur année de formation. Pour les étudiants de deuxième année, il s'agit principalement d'un travail de type « séminaire ou exposé » destiné à évaluer la maîtrise des contenus de cours et de la rédaction d'un texte « scientifique » par la mobilisation de compétences langagières et de recherche documentaire.

*[...] c'est un dossier qu'on doit rendre par rapport à une thématique qu'on développe et il y a aussi toujours la question des sources qui est nouvelle pour nous aussi. [2FM<sup>2</sup>]*

En troisième année, les étudiants voient ces tâches comme le traitement d'une situation plus ou moins authentique en utilisant des acquisitions de la formation. Il s'agit alors de synthétiser et analyser les apprentissages faits en cours et sur le terrain, autrement dit de manifester sa compétence à faire des liens théorie-pratique et à réfléchir sur sa pratique.

*Je dirais que c'est une sorte d'exposé, de travail à rendre et qui est différent de ce qu'on connaît du collège parce que ça englobe différents points de vue et il y a toujours un problème derrière, c'est pas quelque chose qu'on peut résoudre facilement... il faut toujours voir différents points de vue. [3QB]*

*C'est souvent des situations qu'on a vécues en stage et puis que à partir de là, on doit faire des liens théorie-pratique donc voir la théorie qu'on a vue en cours et la mettre en lien avec la situation qu'on a vécue... [3BD]*

*C'est un travail de recherche dans tous les documents qu'on a vus en cours... un travail de synthèse des différents cours... aller piocher un peu partout des informations et puis les appliquer à une tâche plus ou moins concrète... ça dépend des tâches complexes, il y en a qui sont plus concrètes que d'autres... [3GF]*

#### ***Perception des critères par les étudiants***

Quelle que soit l'année de formation, les étudiants n'ont qu'une connaissance partielle des critères utilisés par les formateurs pour corriger leurs tâches. Si la plupart d'entre eux sont capables de dire que plusieurs critères sont utilisés, ils ne parviennent toutefois pas à les citer avec précision.

*Y'a une feuille à la fin avec les 3+ les 2+ les - c'est les seuls critères qu'on a. Mais après ce qu'il y a dedans et ce qui est demandé, c'est un petit peu abstrait, ça change je pense suivant les profs. [2WI]*

*Bien y'a une base de 4 critères je crois dans chaque critère y'a des sous points et puis ça prend en compte le contenu et la forme et puis voilà quoi. [...] Les critères, ouf je me souviens plus, y'a la forme où dedans ça comprend l'orthographe, la notation des sources entre autres. Dans le contenu, je crois qu'il y a la cohérence et y'a aussi les ..., la pertinence des informations, si c'est en lien avec le thème qui était demandé puis sinon ... le contenu ou la forme ... [2BG]*

Afin d'assurer une cohérence tout au long de la formation, les critères ont été voulus identiques dans toutes les tâches. Cette similitude est plus ou moins identifiée par les étudiants, bien qu'ils aient déjà effectué au minimum trois tâches. Quelques étudiants affirment avec certitude que les critères sont les mêmes pour toutes les tâches. D'autres pensent que certains éléments sont semblables, tandis que certains varient.

*Je trouve que oui, grosso modo le tableau d'évaluation c'est toujours un peu le même, c'est enfin surtout au niveau de la forme c'est toujours le même puis après au niveau du contenu, ça change un peu au niveau de la TC, mais à la base, c'est toujours les mêmes exigences. [3GF]*

Enfin, pour quelques étudiants, les critères sont clairement différents selon la nature des tâches.

*Non, je pense pas. On a eu des TC où on devait créer des films, donc là il y a eu tout ce qui était plus média qui rentrait en ligne de compte. D'autres où c'est plus sur des situations concrètes, ou alors sur vraiment une thématique plus large. Parfois on a des choix, d'autres pas donc ça change souvent. [3MB]*

Quand il s'agit de préciser ce que recouvre chacun des critères, les étudiants proposent spontanément une définition de la forme dans laquelle on retrouve systématiquement ce qui a trait aux normes de présentation des références et citations, mais aussi à la correction langagière. Par contre, pour les autres critères, ils éprouvent de grandes difficultés à les définir et des divergences apparaissent. Globalement, ils parlent d'un critère « contenu » dans lequel on trouve mentionné de manière peu distincte la pertinence, la cohérence et plus rarement la qualité. Chez les étudiants de deuxième année, le critère « contenu » englobe essentiellement le recours adéquat à des concepts théoriques vus dans les cours et leur compréhension.

*Y'avait le thème [...] du contenu travaillé si c'était sur la pédagogie, est-ce que on est dans le thème est-ce qu'on a bien compris le concept du triangle pédagogique ou comme ça. [2SE]*

Les étudiants de troisième année reprennent ces aspects tout en y ajoutant l'intérêt et la cohérence du travail ainsi que des éléments d'analyse, en particulier la capacité à faire des liens théorie-pratique.

*La pertinence de l'analyse, la pertinence des liens qui sont faits entre la théorie et la pratique, l'utilisation à bon escient des ressources et puis voilà. Les liens théorie pratique, ça j'ai déjà dit quoi ... et que ça reste dans le thème et que y'aille de l'intérêt dans le travail. [3FC]*

Si tous pensent que le contenu est plus important que la forme, les étudiants de deuxième année ont l'impression que parfois ce dernier critère compte pour une part très importante dans l'évaluation de leur travail.

*Je pense qu'il y en a certains qui valent plus que d'autres, dans la cohérence enfin je pense que le contenu il est plus important que la forme. Je pense que s'il y a quelques fautes d'orthographe, je pense que c'est moins important que si on loupe complètement, si on est hors sujet. Mais j'ai l'impression qu'ils axent quand même beaucoup en tout cas en 1<sup>e</sup> sur la partie moi je dis ça comme ça, mais démarche scientifique pour les sources ou comme ça. Comme c'est nouveau, ils veulent nous reprendre directement pour qu'on apprenne juste et pas qu'en troisième on se soit habitué à faire des erreurs que ce soit vraiment un problème en troisième parce qu'en fait on fait faux depuis trois ans. J'ai l'impression que cette partie là, ils y mettent quand même de l'importance. [2SE]*

Pour les étudiants plus avancés dans leur formation, l'importance de la forme est nuancée. En effet, ils estiment généralement qu'étant donné leur futur métier, la forme est importante puisqu'un enseignant doit être capable de produire des textes correctement rédigés et présentés ; toutefois, la forme doit rester subordonnée aux autres critères, et ne pas justifier à elle seule un échec sur la tâche.

*Ils ont tous leur importance, non... enfin. La dernière TC on a pu louper pour la forme mais pas tout refaire, mais à mon avis ça compte aussi beaucoup, parce qu'on est des futurs enseignants donc y faudrait quand même savoir écrire un minimum juste donc pour moi y sont tous importants avec peut-être un petit peu moins sur la forme parce que ça permet de corriger sans tout recommencer, mais... [3FC]*

### **Indépendance des critères**

Concernant l'indépendance des critères, les étudiants sont plutôt sceptiques ; ils estiment qu'il est peu vraisemblable d'être en échec sur un critère et de réussir les autres, à l'exception peut-être de la forme.

*Si on n'a pas déjà compris la consigne de base c'est clair qu'après on sera totalement faux sur le processus à avoir, sur le déroulement, sur la démarche pour faire une TC. On peut bien réussir la forme pis au niveau de la pertinence de la qualité on a zéro point [...] ceux là sont plus liés... [3BD]*

L'indépendance des critères est une condition importante pour garantir une évaluation valide. Les quatre critères choisis pour évaluer les TC étant censés respecter cette condition, les auteurs ont testé leur indépendance en comparant les appréciations données par les évaluateurs à chacun des critères pour les deux tâches. Une seule valeur étant associée à la non-acquisition, les auteurs ont opté pour le test du khi-deux, mieux à même de traiter l'association entre les appréciations.

Sur l'ensemble des tâches, les auteurs constatent une dépendance assez importante entre tous les critères, tous les coefficients khi-deux s'avérant significatifs. Autrement dit, dans une même tâche, les appréciations des évaluateurs sur les critères de pertinence, cohérence et qualité de la tâche sont liées entre elles. Le critère forme – bien que les coefficients soient un peu plus faibles – est évalué en concordance avec les autres critères.

Tableau 3  
***Indépendance des critères pour les tâches complexes acquises et non acquises***

	TC réussies (n=106)	Pertinence	Cohérence	Qualité dev.	Qualité analyse
	TC échouées (n=52)	Khi2 (ddl)	Khi2 (ddl)	Khi2 (ddl)	Khi2 (ddl)
Cohérence	TC réussies	52,6 (4)**			
	TC échouées	10,7 (9)			
Qualité dev.	TC réussies	24,3 (6)**	33,4 (6)**		
	TC échouées	5,8 (9)	20,7 (9)*		
Qualité analyse	TC réussies	5,4 (4)	15,8 (4)**	9,7 (4)*	
	TC échouées	10,7 (6)	3,7 (6)	23,2 (4)**	
Forme	TC réussie	21,5 (6)**	17,9 (6)**	96,5 (9)**	11,1 (6)
	TC échouées	8,6 (9)	12,0 (9)	17,7 (9)*	10,4 (6)

\*\*= $p < .01$  \*= $p < .05$

La distinction entre les tâches réussies et échouées nuance toutefois ces résultats de façon intéressante. En effet, si on considère les tâches acquises, cette interdépendance de l'ensemble des critères est bien présente. Par contre, celle-ci disparaît presque systématiquement lorsque seules les tâches échouées sont prises en compte (tableau 3). Les analyses réalisées sur ces tâches



montrent une absence quasi totale de lien dans la manière d'évaluer les différents critères. Tout se passe comme si, face à une tâche qu'ils estiment suffisante, les évaluateurs se basent sur leur impression globale pour apprécier les différents critères, leur attribuant des valeurs semblables ou proches. On est en présence d'un fort effet de contamination : si la tâche est bonne, alors elle est pertinente, cohérente et de qualité et sa forme est correcte. Par contre, face à des tâches non acquises, les évaluateurs nuancent leurs jugements et notent de manière différenciée les quatre critères. La forme peut être mauvaise dans un travail cohérent et pertinent, alors qu'un travail cohérent peut ne pas être considéré comme pertinent ou de qualité.

Une observation plus détaillée des causes des échecs – un seul critère noté insuffisant justifiant la non-acquisition – illustre cette indépendance des critères. Ainsi, sur l'ensemble des TC1B et TC2E échouées (n = 52), 14 le sont uniquement pour des aspects formels.

### ***Tâche, étudiant, formateur : trois pôles d'un système***

Dans le cadre de la section précédente, en se basant sur les données faisant référence principalement aux TC1B et TC2E, les auteurs ont démontré que les étudiants avaient développé une compréhension peu définie et très lacunaire des tâches complexes et des différents critères d'évaluation. Au-delà de ce premier niveau d'analyse, le discours des étudiants sur l'évaluation par TC en général a amené un nouveau questionnement sur les liens unissant les trois éléments principaux du dispositif : la tâche, le formateur et l'étudiant (voir figure 1).

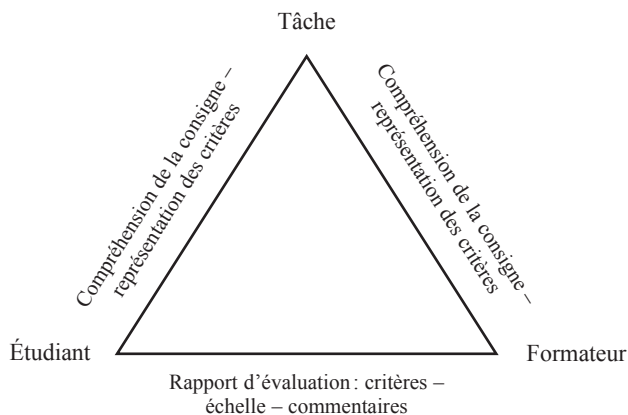


Figure 1. *Les trois pôles de l'évaluation par tâche complexe*

### *Liens étudiant et tâche : consigne et critères*

À ce niveau, l'analyse des propos des étudiants met en évidence plusieurs éléments : premièrement, des représentations individuelles diverses, tant en ce qui concerne les quatre critères utilisés que l'évaluation en général. Deuxièmement, si les critères sont souvent compris de manière différente, il semble que les étudiants entretiennent globalement le même rapport à la consigne dans le sens où celle-ci les accompagne durant tout le processus de réalisation de la tâche.

*[Comment faites-vous au moment de réaliser la tâche?] J'ai la feuille de consignes à côté et j'écris d'abord les parties, je fais d'abord la table des matières et puis après je remplis... [et vous faites la table des matières avec la consigne?] Je prends exactement ce qu'il y a dessus... [3OD]*

Puisque, pour les étudiants, la consigne de la tâche est perçue comme l'élément central de la réalisation, elle va donner lieu à des discussions entre eux afin d'essayer de clarifier les zones d'ombre et de confronter les diverses compréhensions possibles.

*[Est-ce que vous échangez avec vos collègues?] Oui, surtout par rapport à la consigne, pour savoir qu'est-ce qu'elles ont mis où... est-ce qu'il faut mettre les éléments dans l'ordre, où est-ce qu'on a une marge de manœuvre... [3BL]*

*A chaque fois que j'écrivais un petit bout, je relisais plusieurs fois la consigne pour voir si tout ce que j'avais écrit était vraiment en lien et puis après en discutant avec les autres de la HEP on a beaucoup discuté pour être sûr qu'on n'était pas à côté... [2BG]*

Les critères d'évaluation, et leur compréhension par les étudiants, sont en fait extrêmement peu utilisés par ces derniers dans le cadre de la réalisation de leur tâche complexe. Ils apparaissent parfois, dans le discours des étudiants, au moment de la vérification finale avant de remettre le travail.

*Alors j'imprime toujours une feuille d'évaluation, alors cette feuille-là [les critères] et aussi la consigne... et puis quand je rédige le travail, j'essaie d'avoir toujours ces documents papier à portée de mains et de m'y référer pour voir si c'est en adéquation ou pas... surtout avec la consigne et une fois que j'ai fini le travail, je me dis alors ok, la forme, c'est bien ce qu'il y a dedans, être sûre que c'est bon...[3BL]*

### *Lien étudiant et formateur: le rapport d'évaluation*

Le lien entre le pôle du formateur, évaluateur de la tâche, et l'étudiant existe essentiellement dans le rapport d'évaluation qui consiste en une appréciation des quatre critères sur une échelle à quatre niveaux (-, +, ++, +++) accompagnée parfois de commentaires à propos de chaque critère. À l'exception des TC de troisième année, les rapports certificatifs ne contiennent pas de mentions, seuls la réussite ou l'échec sont indiqués.

L'analyse des réponses des étudiants à propos de cette interface entre le formateur et eux-mêmes montre que, dans un premier temps, c'est bien le résultat de la tâche qui prime, avant la prise de connaissance de la réussite des différents critères. Il est intéressant de noter que, malgré une culture importante de l'évaluation chiffrée, les étudiants ne cherchent pas à « transformer » le niveau d'acquisition des différents critères en une note.

Les auteurs ont également mis en évidence l'importance que les étudiants accordent aux commentaires des formateurs. Ces éléments qualitatifs sont considérés par les étudiants comme un juste retour sur l'investissement en temps consacré à la réalisation de la tâche.

*Les commentaires, ça aide à comprendre la croix qui est mise... sinon ça fait un peu, voilà elle a fait sa tâche complexe et puis c'est bon... C'est quand même des choses dans lesquelles je vais à fond dans ces tâches complexes, et j'aime bien, c'est l'occasion de creuser un sujet, alors je vais chercher, je vais lire... même si ça m'emôte de devoir écrire et tout ça... ça m'intéresse donc si on me pose des questions, c'est que ça a intéressé la personne qui a lu... [3OD]*

Une absence de commentaire est même considérée comme frustrante.

*J'étais contente de mon évaluation, mais j'étais un peu déçue, il n'y avait aucun commentaire... c'est un peu frustrant... [2BG]*

Les étudiants mettent assez clairement en évidence deux types de commentaires : ceux jugés non pertinents et ceux considérés comme pertinents.

Les premiers, considérés comme non pertinents, sont centrés sur le produit et ne font que corroborer les indications fournies par l'échelle d'évaluation du critère.

*Je ne reviens pas sur les commentaires si les remarques sont uniquement sur la tâche, mais si elles peuvent m'être utiles parce qu'elles sont plus générales, oui... [3BL]*

*Là par exemple, elle dit « les aspects formels sont corrects », super mais ça ne m'apporte rien à moi... si déjà elle met trois +, ben je comprends... [2TH]*

*Par exemple, ceux-là ils ne m'apportent rien, ils ne mettent pas en question, ils ne me font pas aller plus loin... je me dis juste « bon ben d'accord... » [3MT]*

Les commentaires jugés pertinents, même en cas de réussite du critère, ne sont pas uniquement centrés sur le produit ou sur la tâche réalisée, mais ils apportent des éléments et des perspectives de régulations à l'étudiant quant au processus.

*Si ils sont pertinents, ça nous questionne quand même... Parfois, il y a un conseil, ou des questions qui peuvent nous relancer... ah, j'avais pas pensé à ça... ou ah, ça j'aurais pu le mettre effectivement... [3MT]*

*Si, au contraire, mais des commentaires constructifs, parce que si il y a trois +, alors c'est déjà que les aspects formels sont corrects... [...] et même s'il y a trois +, il y a toujours des choses qui peuvent être refaites... Avec les commentaires, on peut mieux se construire pour la prochaine tâche par exemple... [2TH]*

*La remarque, quand elle est fondée et puis qu'on la comprend, c'est celle qui apporte le plus au travail... Parce qu'on peut revenir là-dessus... [3MB]*

Le commentaire jugé pertinent permet même à l'étudiant d'accepter une part importante de la subjectivité de l'évaluation.

*[Sur ce travail en particulier (l'étudiant a eu +++) si vous aviez eu +, vous auriez eu l'impression que c'est évalué de manière juste?] Ben, il faudrait voir les commentaires pour savoir qu'est-ce qui était moins bien... après, si c'est expliqué pourquoi il n'y a qu'un +, ben ça me va, il faut juste une bonne justification et pas uniquement, par exemple, « manque de pertinence »... [2BG]*

### **Lien formateur et tâche : les fonctions de l'évaluation**

La question des représentations des formateurs à propos des critères d'évaluation et des consignes des tâches complexes n'a pas été abordée dans la présente étude. Néanmoins, la connaissance du contexte amène à penser que la culture commune à ce sujet pourrait être également étudiée. Par contre, cette recherche permet de distinguer un élément intéressant qui concerne les diverses fonctions de l'évaluation attribuées à ce dispositif par les formateurs. En effet, les auteurs l'ont illustré précédemment, les commentaires du rapport d'évaluation sont jugés par les étudiants comme étant plus ou moins pertinents. En analysant des rapports d'évaluation produits par les formateurs, les auteurs ont distingué deux types d'évaluateur : les uns « formant » et les autres « certifiant ». Les premiers semblent profiter de l'analyse de la tâche de l'étudiant pour fournir à celui-ci des éléments de formation. Dans ce cas, les commentaires sont souvent étoffés, précis et sont centrés sur des ressources et leur mobilisation dans le cadre du processus de production de la tâche. Les seconds

semblent accorder à l'évaluation de la tâche une unique fonction de certification de l'acquisition des ressources et de leur mobilisation. Dans ce cas, les commentaires sont soit inexistants, soit ils ne font que reprendre les indicateurs des critères.

Les fonctions de l'évaluation (Allal, Cardinet, & Perrenoud, 1988 ; Cardinet, 1988) devraient être questionnées à la lumière de ces éléments. En effet, sur le plan de l'institution, ce dispositif d'évaluation s'inscrit bien dans une perspective certificative, mais quant aux étudiants, l'investissement en temps et le sens qu'ils construisent autour de ces tâches créent une attente importante quant au retour des formateurs. Un questionnement fait jour : les fonctions classiques et conceptuellement imperméables de l'évaluation sont-elles toujours pertinentes dans un type d'évaluation comme celui-ci ?

### *Construction du sens de l'évaluation par tâche complexe*

Finalement, les entretiens réalisés et leur analyse ont donné l'occasion de mettre en évidence le sens construit par les évalués à propos de ce mode d'évaluation.

Ce qui ressort de manière générale, c'est que les étudiants disent préférer une évaluation par tâche complexe à une évaluation plus «classique» à la manière des examens sur table. Cette préférence, affirmée par 15 étudiants sur les 16 interrogés, est essentiellement justifiée par la question de l'apprentissage en général et en particulier par le transfert des ressources en situations réelles inhérent à ce type de tâche.

*... je me lance un challenge et j'essaie de faire la TC aussi sur ça parce que je trouvais aussi plus intéressant au niveau de l'analyse de faire sur quelque chose que je ne maîtrise pas à 100%... [3GF]*

*L'analyse de situation ça nous oblige à prendre plusieurs points de vue et à intégrer ce qui a été vu dans les cours théoriques... [3OD]*

*J'ai pu tester la démarche inductive parce que je ne l'avais jamais fait jusqu'à maintenant... donc moi j'ai appris des choses donc c'est une chose que je referai sûrement quand j'aurai ma classe... [3BD]*

*A contrario*, l'évaluation «classique», bien connue des étudiants, est perçue de manière beaucoup plus négative, essentiellement en raison de son aspect peu authentique imposant souvent une centration sur des contenus dans une temporalité très courte.

*Parce que finalement c'est recracher des connaissances qu'on doit apprendre, enfin mémoriser et puis finalement [...] je pense que 90% des gens c'est la même*

*chose, ce genre de connaissances on les mémorise, on a l'examen, pis trois heures après c'est formaté... [3FC]*

*L'examen c'est très sélectif, c'est vraiment sur un petit truc, si on sait peut-être pas un deux trucs on a mal répété ou qu'on a pas trop compris ben y'a des chances qu'on loupe là... [3QB]*

Globalement, le mode d'évaluation par TC semble donc avoir du sens pour les étudiants parce qu'il leur permet d'établir des liens entre ce qui leur est demandé et la réalité de leur future profession. Afin de mieux comprendre l'influence de la tâche en tant que telle sur la construction du sens de l'évaluation, les étudiants ont été invités à s'exprimer sur l'intérêt de chaque tâche complexe utilisée à la HEP FR. Le schéma ci-après (figure 2), met en évidence, à partir de leur discours, les caractéristiques des diverses tâches et le sens qui leur est attribué.

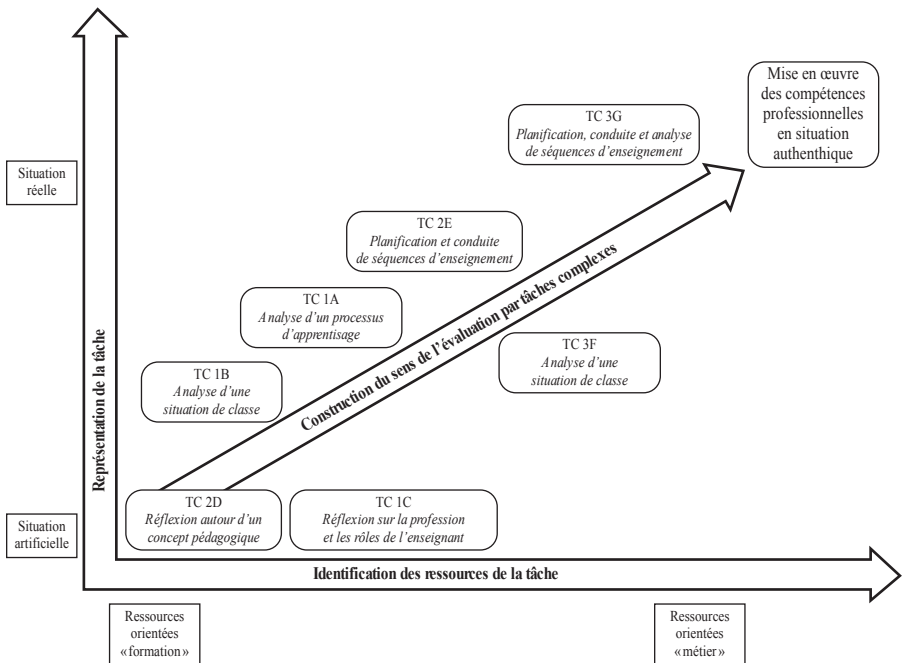


Figure 2. *Sens attribué aux tâches complexes et à leur évaluation*

En abscisse se trouve l'identification des ressources mobilisées dans le cadre de la tâche. Celles-ci sont perçues par les étudiants comme étant orientées « formation », c'est-à-dire considérées comme nécessaires à la réussite de leurs études, ou orientées « métier », c'est-à-dire nécessaires à l'exercice de la profession d'enseignant.

En ordonnée, la représentation de la tâche correspond à l'identification de la nature artificielle ou réelle de la situation ; autrement dit, la tâche est-elle perçue comme pouvant être transposée directement ou non à la pratique professionnelle ? Chaque tâche est ensuite placée dans ce système d'axes. Il apparaît que la construction globale du sens de l'évaluation par tâche complexe est également liée au parcours des étudiants sur leurs trois années de formation, les premières tâches rencontrées étant moins porteuses de sens parce que mobilisant des ressources plutôt orientées « formation ».

Il est possible de bien identifier dans le discours des étudiants que si, de manière générale, les TC sont porteuses de sens et d'apprentissage, celles qui se centrent essentiellement sur des ressources développées et attendues en formation semblent avoir moins de sens pour eux :

*[Les TC 1A (Analyse d'un processus d'apprentissage) et 1B (Analyse d'une situation de classe) sont-elles pertinentes pour la pratique ?] Peut-être pas directement, mais dans le fond oui parce que c'est quand même par ces tâches qu'on se replonge dans les cours pis qu'on comprend mieux les mieux les concepts pis qu'après les concepts sont applicables dans la classe... [2SE]*

*Je pense que celle au début de cette année [2D Réflexion autour d'un concept pédagogique] [...] on devait analyser la phrase d'un pédagogue [...] c'est bien pour le futur, mais après je sais pas ce qui va nous aider plus tard, c'est vraiment un concept. On pourra peut-être l'utiliser, on verra... [3QB]*

*Le problème c'est que une fois que c'est fini, je ne sais pas si on va s'en rappeler trois ans plus tard quoi... en flou on aura des souvenirs de ce que ça nous aura appris... mais après ça nous apprend aussi beaucoup au niveau de comment écrire un texte, comment structurer, comment analyser... [3GF]*

Par contre, les tâches qui demandent la mobilisation de ressources directement en lien avec la pratique ou une posture à développer en tant qu'enseignant sont réellement porteuses de sens pour les étudiants :

*Oui, on doit faire ça tout le temps planifier dans la continuité [2E Planification et conduite de séquences d'enseignement-apprentissage] et puis revenir sur ce qu'on a fait dans sa leçon ou ses leçons... pour voir où ils en sont, qui c'est qu'on a touché ou pas, qu'est-ce qu'on devrait faire autrement, qu'est-ce qui a bien marché ou pas... Je pense que oui, c'est un peu la base du futur métier... [3MB]*

*Je trouve qu'on peut vraiment bien relier tout ce qui est cours avec ce qu'on voyait sur la vidéo, on fait des liens. J'ai pas trouvé que c'est une tâche très difficile. On sera confronté à ça je pense en tout cas ben ces situations où ça arrive souvent ce genre. Ouais ça peut arriver en cours... [3QB]*

*Pour moi elle a vraiment du sens [1B] ... c'était peut-être celle qui avait le plus de sens... Parce que j'ai été contraint pour répondre aux attentes de plonger dans certaines lectures pour élargir mon champ de vision... [2KQ]*

Il apparaît donc globalement que plus l'étudiant avance dans son parcours, plus les tâches complexes font appel à des ressources qu'il identifie comme nécessaires à l'exercice de sa profession et intégrées à de la pratique en stage. Le sens de ce mode d'évaluation est alors construit essentiellement parce qu'il demande l'expression de compétences professionnelles dans des situations «réelles» ou proches d'une pratique future.

Pour terminer cette analyse du sens attribué à l'évaluation par tâche complexe, la problématique du caractère formateur de ce mode d'évaluation, déjà identifié lors de l'analyse des commentaires des évaluateurs peut également être mise en évidence. Les apprentissages effectués dans le cadre de la réalisation de ces tâches et relevés par les étudiants mettent bien en évidence l'existence d'une perméabilité, conceptuellement questionnante, entre les fonctions formative et certificative.

*Oui on apprend quelque chose... on apprend oui... c'est fait pour ça... je pense... [Et dans une évaluation classique, on apprend quelque chose?] Non parce que dans une évaluation classique, on redonne ce que l'on a appris, sans l'avoir en fait utiliser dans un contexte, alors que dans une TC, on apprend à faire des liens entre les différentes choses et puis on apprend aussi à réfléchir sur quelque chose qu'on a vu... Ça va plus loin... [3GF]*



## Conclusion

Menée dans une visée exploratoire, cette recherche montre, dans un contexte particulier, le rapport entretenu par les étudiants avec une évaluation des compétences professionnelles en développement au moyen de tâches complexes. Le caractère situé des modalités d'évaluation investiguées et la perspective ancrée de l'approche méthodologique orientent la discussion vers un questionnement du dispositif mis en place, mais aussi sur des aspects plus généraux de l'évaluation des compétences.

La recherche met en évidence le fait que les étudiants n'ont qu'une connaissance partielle et lacunaire du dispositif d'évaluation mis en place. Ils peinent à définir précisément les caractéristiques d'une telle évaluation de même que les critères utilisés, bien que ceux-ci soient identiques pour toutes les tâches. De fait, les étudiants se centrent essentiellement sur la consigne de la tâche pour élaborer leur production.

Au cœur de l'analyse, deux problématiques liées entre elles ont émergé. Elles concernent, d'une part, la nature des tâches, et, d'autre part, le sens qui leur est attribué.

En effet, il est surprenant de constater que malgré un caractère certificatif institutionnellement établi, les tâches sont perçues par les étudiants comme possédant un fort potentiel d'apprentissage. Les régulations qu'ils sont amenés à envisager en fonction des commentaires des formateurs accroissent la perception d'un caractère formatif de telles tâches. Cet aspect est également amplifié puisque les étudiants comparent les tâches complexes à d'autres formes d'évaluation auxquelles ils ont eu l'habitude d'être confrontés comme les examens sur table. Cette ambiguïté entre les fonctions sommative et formative interpelle fortement. Deux interprétations s'offrent. Soit il faut considérer le dispositif mis en place comme manquant de validité, puisque les tâches semblent évaluer bien plus que ce qui a été appris, soit il s'avère nécessaire de considérer que les deux fonctions ne sont pas si hermétiques, le caractère sommatif de telles tâches n'empêchant pas à la fois des apprentissages et des régulations pour les étudiants. Dans le premier cas, une piste pour accroître la validité pourrait être de proposer aux étudiants des tâches isomorphes (Jonnaert, 2011) dans une perspective formative tout en étant conscient qu'une telle procédure risque de s'avérer chronophage. Dans le deuxième cas, c'est

le concept même d'évaluation par tâche complexe qui est à repenser à l'aune des questions suivantes : la résolution d'une situation complexe inédite peut-elle être exempte d'apprentissages ? Et, si des apprentissages devaient avoir lieu, quel statut faudrait-il leur accorder dans une perspective d'évaluation certificative ?

À propos du sens construit par les étudiants, il est intéressant de noter que même en cas d'échec ou de réserves sur l'objectivité du jugement des évaluateurs, les étudiants plébiscitent ce mode d'évaluation par rapport à des formes plus conventionnelles. Autre élément surprenant, contrairement à ce qu'on pourrait penser, le sens des tâches ne se construit pas uniquement à partir de l'authenticité des situations, mais également en fonction de l'identification des ressources considérées comme « utilisables » dans leur pratique future parce que souvent procédurales (Perrenoud, 1996) et, dans une moindre mesure, mobilisables dans leur formation.

En conclusion, la mise en évidence du sens construit par les étudiants dans le cadre de l'évaluation de leurs compétences témoigne de l'intérêt du dispositif mis en place à la HEP-FR. Toutefois, les résultats montrent aussi qu'un dispositif d'évaluation par TC s'accompagne d'une grande complexité et nécessite des ajustements. Pour accroître sa validité, les pistes peuvent s'orienter sur différents éléments qui s'inscrivent dans les propositions décrites par Scallon (2004). Ainsi, pour éviter la production d'apprentissages nouveaux pour réussir les tâches, il faudrait veiller à une plus grande cohérence entre les différents dispositifs de formation afin d'offrir aux étudiants davantage de situations d'intégration des ressources développées. Dans le but d'augmenter encore le sens des tâches proposées, une analyse et une restructuration de certaines d'entre elles jugées moins pertinentes par les étudiants s'avèreraient judicieuses. Dans ce cas, une réflexion sur la nature des ressources et le caractère réel des situations proposées permettrait un gain qualitatif certain. Dans cette perspective, il semblerait pertinent de mettre en lien, dans un travail de recherche futur, l'identification par les étudiants des compétences nécessaires à la réalisation des tâches (Luisoni & Monnard, 2012) avec celles déclarées par les concepteurs et celles identifiées par les enseignants du terrain (Coen et al., 2008). Finalement, pour garantir une évaluation véritablement instituée, un travail de coconstruction autour des critères est nécessaire tant chez les formateurs-évaluateurs que chez les étudiants.

## NOTES

1. Les ateliers sont des journées de travail animées par plusieurs formateurs, dont au moins un formateur praticien enseignant primaire. Les problématiques traitées sont souvent interdisciplinaires et visent à renforcer la pratique réflexive des étudiants autour des liens théorie-pratique.
2. Chaque code correspond à un étudiant. Le chiffre indique l'année de formation (deuxième année ou troisième année) et les lettres, des initiales correspondant à un nom fictif.

## RÉFÉRENCES

- Allal, L. (2002). Acquisition et évaluation des compétences en situation scolaire. In J. Dolz & E. Ollagnier (Eds.), *L'énigme de la compétence en éducation* (p. 77-94). Bruxelles, Belgique : De Boeck Université.
- Allal, L., Cardinet, J., & Perrenoud, P. (Eds.). (1989). *L'Évaluation formative dans un enseignement différencié : actes du colloque à l'Université de Genève, mars 1978* (5<sup>e</sup> éd. rev. et corr.). Berne, Suisse : P. Lang.
- Aubret, J., & Gilbert, P. (2003). *L'évaluation des compétences*. Sprimont, Belgique : Mardaga.
- Barbier, J.-M. (1985). *L'évaluation en formation*. Paris, France : PUF.
- Beckers, J. (2002). *Développer et évaluer des compétences à l'école : vers plus d'efficacité et d'équité*. Bruxelles, Belgique : Éditions Labor.
- Cardinet, J. (1988). *Pour apprécier le travail des élèves*. Bruxelles, Belgique : De Boeck-Wesmael.
- Coen, P.-F., Galland, A., Monnard, I., Perrin, N., & Rouiller, Y. (2008). Pertinence et validité des tâches complexes dans l'évaluation de futurs enseignants. In G. Baillat, J.-M. De Ketele, L. Paquay, & C. Thélot (Eds.), *Évaluer pour former* (p. 159-175). Bruxelles, Belgique : De Boeck Université.
- Crahay, M. (2006). Dangers, incertitudes et incomplétude de la logique de la compétence en éducation. *Revue française de pédagogie*, 154, 97-111.
- De Ketele, J.-M., & Gérard, F.-M. (2005). La validation des épreuves d'évaluation selon l'approche par les compétences. *Mesure et évaluation en éducation*, 28(3), 1-26.
- Dolz, J., & Ollagnier, E. (Eds.). (2002). *L'énigme de la compétence en éducation*. Bruxelles, Belgique : De Boeck Université.
- Galland, A., Monnard, I., & Coen, P.-F. (2010). Évolution de tâches complexes destinées à évaluer des enseignants en formation dans deux institutions de Suisse romande. *Revue des HEP et institutions assimilées de Suisse romande et du Tessin*, 11, 145-160.
- Gérard, F.-M. (2009). *Évaluer des compétences : guide pratique* (2<sup>e</sup> éd.). Bruxelles, Belgique : De Boeck Supérieur.
- Hadji, C. (2012). *Faut-il avoir peur de l'évaluation ?* Bruxelles, Belgique : De Boeck.

- Jonnaert, P. (2011). Sur quels objets évaluer des compétences? *Éducation & Formation, e-296*, 31-43. Retrieved from <http://ute3.umh.ac.be/revues/include/download.php?idRevue=12&idRes=106>
- Le Boterf, G. L. (2006). *Ingénierie et évaluation des compétences* (5<sup>e</sup> éd.). Paris, France : Éditions d'Organisation.
- Luisoni, M., & Monnard, I. (2012, Janvier). *Quel sens de futurs enseignants attribuent-ils à l'évaluation par tâche complexe?* Communication présentée au 24<sup>e</sup> colloque de l'ADMEE-Europe Luxembourg, Luxembourg. Abstract retrieved from <http://admee2012.uni.lu/pdfdownload.html#5>
- Miles, M., & Huberman, M. (2003). *Analyse des données qualitatives* (2<sup>e</sup> éd.). Paris, France : De Boeck Université.
- Mottier Lopez, L. (2009). La reconnaissance d'une professionnalité : un processus situé d'évaluation interprétative. In A. Jorro (Ed.), *La reconnaissance professionnelle : évaluer, valoriser, légitimer* (p. 123-145). Ottawa, Canada : Presses de l'Université d'Ottawa.
- Perrenoud, P. (1996). Savoirs de référence, savoirs pratiques en formation des enseignants : une opposition discutable. *Éducation et Recherche, 2*, 234-250.
- Perrenoud, P. (1999). *Dix nouvelles compétences pour enseigner*. Paris, France : ESF.
- Rey, B. (2008). Quelques aspects éthiques de l'évaluation. In C. Thélot, J.-M. De Ketele, & G. Baillat (Eds.), *Évaluer pour former* (p. 57-67). Bruxelles, Belgique : De Boeck Université.
- Roegiers, X. (1999). Savoirs, capacités et compétences à l'école : une quête de sens. *Forum-pédagogies, mars*, 24-31.
- Roegiers, X. (2003). *Des situations pour intégrer les acquis des élèves*. Paris, France : De Boeck Université.
- Roegiers, X. (2004). *L'école et l'évaluation : des situations pour évaluer les compétences des élèves*. Bruxelles, Belgique : De Boeck Université.
- Roegiers, X., & De Ketele, J.-M. (2000). *Une pédagogie de l'intégration : compétences et intégration des acquis dans l'enseignement*. Bruxelles, Belgique : De Boeck Université.
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Bruxelles, Belgique : De Boeck Supérieur.
- Strauss, A., & Corbin, J. (2004). *Les fondements de la recherche qualitative : techniques et procédures de développement de la théorie enracinée*. Fribourg, Suisse : Academic Press Fribourg.
- Tardif, J. (2006). *L'évaluation des compétences : documenter le parcours de développement*. Montréal, Canada : Chenelière-éducation.

Date de réception : 5 novembre 2012

Date de réception de la version finale : 14 août 2013

Date d'acceptation : 10 septembre 2013