

Évaluation des compétences dans un programme de formation en enseignement

Validité de construit curriculaire

Léon Harvey

Volume 35, numéro 2, 2012

Date de réception : 7 mars 2012

Date de réception de la version finale : 15 octobre 2012

Date d'acceptation : 17 octobre 2012

URI : <https://id.erudit.org/iderudit/1024721ar>

DOI : <https://doi.org/10.7202/1024721ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (imprimé)

2368-2000 (numérique)

[Découvrir la revue](#)

Citer cet article

Harvey, L. (2012). Évaluation des compétences dans un programme de formation en enseignement : validité de construit curriculaire. *Mesure et évaluation en éducation*, 35(2), 69–95. <https://doi.org/10.7202/1024721ar>

Résumé de l'article

Cette étude propose d'analyser les dossiers académiques, qui cumulent les crédits mérités dans le cadre d'un programme, afin de déterminer si les notes, obtenues durant les cours associés à la formation, permettent de porter un jugement valide sur le développement des compétences attendues. Le concept de validité de construit curriculaire est introduit afin de qualifier le jugement qui sera porté à cet effet. Des analyses factorielles confirmatoires permettent de confronter la structuration Cours-Compétences sous-jacente au programme avec l'empirie, ce qui permet de valider les jugements énoncés. Les analyses portent sur les dossiers académiques de 164 finissantes et finissants du programme d'enseignement secondaire d'une université québécoise, diplômés entre 2002 et 2010. Les résultats confirment que les évaluations se structurent en fonction de quatre groupes de compétences liés aux fondements de l'éducation, à l'acte d'enseigner, aux aspects sociaux et au développement de l'identité professionnelle. Les avantages et les limites de cette analyse sont discutés.

Évaluation des compétences dans un programme de formation en enseignement : validité de construit curriculaire

Léon Harvey

Université du Québec à Rimouski

MOTS CLÉS : évaluation, compétences, validité de construit curriculaire, analyse factorielle confirmatoire, éducation

Cette étude propose d'analyser les dossiers académiques, qui cumulent les crédits mérités dans le cadre d'un programme, afin de déterminer si les notes, obtenues durant les cours associés à la formation, permettent de porter un jugement valide sur le développement des compétences attendues. Le concept de validité de construit curriculaire est introduit afin de qualifier le jugement qui sera porté à cet effet. Des analyses factorielles confirmatoires permettent de confronter la structuration Cours-Compétences sous-jacente au programme avec l'empirie, ce qui permet de valider les jugements énoncés. Les analyses portent sur les dossiers académiques de 164 finissantes et finissants du programme d'enseignement secondaire d'une université québécoise, diplômés entre 2002 et 2010. Les résultats confirment que les évaluations se structurent en fonction de quatre groupes de compétences liés aux fondements de l'éducation, à l'acte d'enseigner, aux aspects sociaux et au développement de l'identité professionnelle. Les avantages et les limites de cette analyse sont discutés.

KEY WORDS: evaluation, competencies, curriculum, confirmatory factor analysis, education

This study proposes to analyse the student's academic records that currently hold the credits gathered from the program framework, to determine if the grades, inherited from the courses, allows a valid judgment about the competencies developed at the end of the curriculum. The concept of curricular construct validity is introduced to qualify the judgment that will be issued. Confirmatory factorial analyses allow confronting the theoretical Course-Competencies mapping against the empirical data in order to validate the judgments issued. The analyses were conducted on 164 students from a university located in the province of Québec, who graduated from a secondary teaching program between 2002 and 2010. The results confirm that the evaluations are structured into four groups of competencies: educational foundations, teaching acts, social aspects and professional identity. Advantages and limits of this analysis will be discussed.

PALAVRAS-CHAVE: avaliação, competências, validade do construído escolar, análise fatorial confirmatória, educação

Este estudo propõe-se analisar os dossiês académicos, que reúnem os créditos adquiridos no quadro de um programa, para determinar se as notas, obtidas durante os cursos associados à formação, permitem formular um juízo válido sobre o desenvolvimento das competências esperadas. O conceito de validade do construído curricular é introduzido para qualificar a formulação deste juízo. As análises fatoriais confirmatórias permitem confrontar a estruturação Curso-Competências subjacentes ao programa com a empiria, o que permite validar os juízos enunciados. As análises incidem sobre os dossiês académicos de 164 finalistas do programa de ensino secundário de uma universidade do Quebec, diplomados entre 2002 e 2010. Os resultados confirmam que as avaliações se estruturam em função de quatro grupos de competências ligados aos fundamentos da educação, ao ato de ensinar, aos aspetos sociais e ao desenvolvimento da identidade profissional. As vantagens e os limites desta análise são discutidas.

Note de l'auteur – Toute correspondance peut être adressée comme suit : Léon Harvey, Unité des sciences de l'éducation de Rimouski, Université du Québec à Rimouski, 300 allée des Ursulines, Rimouski (Québec), G5L 3A1, téléphone : 418 723-1986, poste 1758, ou par courriel à l'adresse suivante : [Leon_Harvey@uqar.ca].

Introduction

Avec l'augmentation du nombre de programmes de formation professionnalisants, il existe des pressions accrues de la communauté scientifique, du public, des médias, des gestionnaires et des acteurs eux-mêmes afin de déterminer comment ces programmes permettent d'attester du développement des compétences (Comité d'agrément des programmes de formation à l'enseignement, CAPFE, 2010; Ryan & Cousins, 2009). Dans certaines professions, la certification professionnelle est réalisée *a posteriori* par des organismes externes autres que l'institution de formation. Ceux-ci sont responsables de la mise en place d'épreuves présentant un bon niveau de validité attestant de la compétence des futurs professionnels. Dans d'autres professions, dont celles relatives à l'éducation, l'admission à l'exercice de la profession repose sur l'obtention par le candidat¹ des crédits prévus au programme de formation (Louis, Jutras, & Hensler, 1996). Dans cette perspective, les programmes de formation doivent démontrer qu'ils prévoient des dispositifs présentant un bon niveau de validité pour la mesure des compétences visées ; mais au final, c'est le dossier académique qui permet d'attester formellement de l'atteinte de l'objectif visé au travers du cumul des différents crédits obtenus. La question fondamentale qui se pose alors est de savoir si ce dossier académique permet de porter un jugement valide afin d'attester du développement des compétences. Afin de qualifier le jugement qui sera produit, le présent article introduira la notion de *validité de construit curriculaire*.

Problématique

Au Québec, la plupart des programmes actuels de formation des enseignants ont été élaborés au milieu des années 90 et ont été révisés au début des années 2000. Les orientations de ces programmes ont été spécifiées à partir du référentiel de compétences proposé par le ministère de l'Éducation du Québec (MEQ, 2001). Celui-ci définit les compétences liées à la profession enseignante alors que les cursus des différentes universités (Gauthier & Mellouki, 2006 ; Harvey & Laroui, 2006) précisent comment les divers cours² d'un programme s'agencent afin de permettre le développement d'un professionnel compétent et cultivé. S'échelonnant sur des baccalauréats de quatre années, les programmes proposent divers cours qui permettent de développer les

compétences liées à la maîtrise des fondements de l'éducation, de l'acte d'enseigner, aux aspects sociaux de la profession ainsi qu'au développement identitaire.

La notion de compétence est au centre de ces programmes. Comme le mentionnent De Ketele et Gérard (2005), «l'approche par compétences cherche à développer la possibilité par les apprenants de mobiliser un ensemble intégré de ressources pour résoudre une situation-problème appartenant à une famille de situations» (p. 2). La compétence est définie comme un savoir-agir complexe qui sous-tend la mobilisation en contexte réel d'une variété de ressources internes et externes (Jonnaert, Barrette, Boufrahi, & Masciotra, 2004; MELS, 2006; MEQ, 2001; Prigent, Bernard, & Kozanitis, 2009). Cette définition suppose qu'une diversité d'éléments doit être mobilisée lors de l'évaluation : des savoirs, des expériences, des attitudes, des habiletés et des ressources. Une des difficultés qu'éprouvent les programmes par compétence est que la mobilisation complète de tous ces éléments ne s'effectue que dans des situations authentiques. Les cours offerts en institution ne permettent que difficilement cette complète mobilisation ; il est reconnu que dans certains cours institutionnels, les étudiantes et étudiants universitaires sont amenés à élaborer des savoirs professionnels (Buisse, 2012) et non des compétences.

En mesure, la compétence est considérée comme un construit latent qui est non directement observable ; la compétence et les savoirs professionnels se construisent tout au long du programme et sont par la suite inférés à partir de la performance dans certaines tâches complexes et d'indicateurs ciblés en situations professionnelles réelles ou simulées. Il existe une vaste documentation (voir Harvey & Barras, 2008) qui suggère que les savoirs acquis en institution peuvent être réinvestis lors de la construction des compétences en situations réelles, en particulier lorsque le milieu professionnel offre un encadrement adéquat (Harvey, 2009) et qu'un langage d'action commun aux deux types d'environnement est partagé (Harvey & Barras, 2008).

Si la documentation sur la notion de compétence et l'approche par compétences est vaste, il existe peu d'études qui permettent de déterminer si ces programmes de formation atteignent bien les objectifs visés. Les évaluations institutionnelles de programme n'offrent qu'une réponse partielle à cette question. Elles relèvent de la politique interne en vigueur dans les universités et ne sont que rarement l'objet d'une diffusion externe. À notre connaissance, au Québec, les rares études externes ont été menées au début des années 2000 par le Bureau du recensement étudiant et de la recherche institutionnelle de

l'Université du Québec (BRERIUQ, 2000a, 2000b) auprès des premiers diplômés et diplômées et de leurs employeurs. Quelques études déplorent que l'organisation des cours dans certains cursus demeure trop morcelée et qu'il existe des variations importantes d'une institution à l'autre quant au poids des composantes psychopédagogique, disciplinaire et de formation pratique (Desjardins & Dezutter, 2009; Gauthier & Mellouki, 2006).

En 2010, le Comité d'agrément des programmes de formation à l'enseignement demande aux universités québécoises de préciser comment elles attestent du développement des compétences professionnelles dans ces programmes. Le comité déplore notamment la disparité des mécanismes d'évaluation des compétences dans les différents cours des programmes. Si les cours associés à la formation pratique mettent en place les dispositifs nécessaires pour l'évaluation des compétences en milieu professionnel, il subsiste des interrogations sur ce qui est effectivement évalué lors des cours se déroulant en contexte institutionnel (CAPFE, 2010).

Ce questionnement sur les évaluations réalisées lors des cours en institution soulève un doute sur l'utilité du dossier universitaire qui, en cumulant les crédits, permet d'attester de l'atteinte du développement des compétences. En soi, le système de notation utilisé n'est pas en cause et est compatible avec l'évaluation des compétences. Au Québec et dans d'autres pays, les universités ont adopté un système de notation littérale où l'évaluation dans le cadre des cours d'un programme de formation prend généralement la forme d'une note unique variant de échec (E) à excellent (A+). Prégent et al. (2009) suggèrent des moyens afin de transformer le jugement sur le développement des compétences en une note finale cadrant avec ce système de notation. Bien que ces auteurs concilient l'évaluation par compétences et l'attribution d'une note littérale, la question de la validité du dossier universitaire comme instrument qui permet d'attester du développement des compétences demeure entière.

En fait, il est primordial, dans un programme par compétences, que le développement et l'évaluation des compétences professionnelles soient planifiés et que le rôle de chacun des cours soit précisément défini en ce sens. D'ailleurs, différents concepteurs de programme ont explicité ce lien à l'aide d'une matrice *Cours-Compétences* (Cardin & Couture, 2006; Harvey & Laroui, 2006; Minier, 2006). Cette matrice synthétise les objectifs de formation à atteindre par chacun des cours et condense toute la logique des programmes. Cependant, cette logique reste théorique et des preuves sont nécessaires afin de démontrer que le programme est efficace et qu'il est adéquatement mis en

œuvre par les ressources professorales qui interviennent dans le programme. Le non-respect de la logique du programme risque d'invalider le processus de développement et d'évaluation des compétences dans les cours et, par conséquent, d'invalider les inférences qui pourraient être faites à partir des notes académiques et du dossier universitaire. Au final, le cumul des crédits prévus au programme ne constitue plus une attestation du développement des compétences. Le présent article propose de vérifier l'adéquation de la matrice théorique *Cours-Compétences* avec l'empirie à partir de l'analyse factorielle confirmatoire afin de déterminer si le dossier universitaire permet d'attester du développement des compétences ou de groupes de compétences professionnelles.

Notre hypothèse s'énonce ainsi et réaffirme que :

H₁. Le dossier universitaire permet de porter un jugement valide sur le développement des compétences ou de groupes de compétences professionnelles.

Cette hypothèse nécessite que la notion de validité soit revisitée. Établir la validité d'un jugement effectué à partir du dossier universitaire, qui cumule les crédits obtenus dans les cours de formation issus d'un cursus défini par compétences, est une opération particulière qui nécessite qu'un nouveau concept soit introduit : *la validité de construit curriculaire*. Pour ce faire, les notions de *validité*, de *validité curriculaire*, de *validité de construit* et de *validité de construit curriculaire* sont présentées.

La *validité* est un concept unitaire. C'est un jugement qui est porté sur l'ensemble des preuves accumulées à propos d'un instrument et qui détermine si les interprétations réalisées à partir de celui-ci sont conformes à l'intention visée (*American Educational Research Association, AERA, 1999*). Plus spécifiquement en éducation, le terme *validité curriculaire* fait référence aux jugements établis par rapport aux objectifs d'un programme « au nom de laquelle les étudiants doivent avoir reçu un enseignement centré sur le contenu et les habiletés nécessaires pour réussir le test » (Scallon, 2004, p. 263). Par exemple, Bodin (2006) remet en question la validité curriculaire d'un test international qui contient environ 25 % de questions qui ne correspondent pas à ce que les élèves français étudient au collège. Dans cette perspective, les évaluations réalisées dans le cadre des cours d'un programme peuvent contenir des éléments qui leurs sont spécifiques et qui ne sont pas prévus au cursus. L'inspection du contenu des épreuves utilisées dans le cadre des évaluations permettrait d'établir cette validité curriculaire, cependant, à l'ordre d'enseignement universitaire, ce contenu est laissé à la discrétion des ressources

professorales et ne peut pas être inspecté sans autorisation. Cet aspect discrétionnaire est une limite importante de la validité curriculaire. Par ailleurs, la validité curriculaire se vérifie cours par cours et cette inspection individuelle ne garantit pas que l'ensemble forme un tout cohérent et conforme à l'intention initiale du programme.

Dans le cadre de la présente analyse, le concept de *validité de construit curriculaire* est introduit. Cette notion se définit comme un jugement qui est porté sur l'ensemble des preuves accumulées à propos d'un programme et qui détermine si les interprétations réalisées à partir de celui-ci sont conformes à l'intention visée, et ce, à partir de l'adéquation entre la structuration théorique des évaluations issues de ce programme avec l'empirie. Ce concept s'inspire à la fois de la notion de *validité de construit* qui évalue l'adéquation entre les structures théorique et empirique (De Ketele & Gérard, 2005 ; Louis et al., 1996) de l'objet mesuré et de celle de *validité curriculaire*.

Le jugement porté sur le programme s'établit à partir d'une lecture intégrée des évaluations qui sont réalisées dans les différents cours qui le constitue, et qui sont inscrites dans les dossiers universitaires. Avec les réserves propres à l'éthique, ces données peuvent être inspectées et analysées dans le cadre d'une recherche.

Le présent article jette donc un regard original et soulève la question de la validité de construit curriculaire lorsque les programmes de formation sont axés sur le développement des compétences. L'étude est pertinente. Ainsi, ce n'est pas parce qu'une étudiante ou un étudiant a réussi l'ensemble des crédits à son programme qu'il a développé les compétences spécifiées par celui-ci. De plus, comme il a été mentionné précédemment, ces crédits proviennent d'activités morcelées ; les évaluations qui sont réalisées n'ont pas forcément été validées ; la validité curriculaire relève de la discrétion des ressources professorales et n'est pas assurée ; et finalement, un jugement global sur la cohérence de l'ensemble des évaluations est nécessaire.

La présente étude vérifie l'adéquation entre les évaluations réalisées dans le cadre des cours avec ce qui est prévu dans le plan de formation. En théorie, s'il existe une certaine cohérence entre ce qui est évalué et ce qui est défini, il sera possible d'affirmer que le curriculum atteint ses objectifs. De même, le dossier universitaire, qui cumule les évaluations, pourra être considéré comme un instrument présentant un bon niveau de validité³.

L'étude de la validité de construit curriculaire permet de cumuler un certain nombre de preuves qui contribuent à établir la validité d'une formation. Elle repose sur le postulat que d'autres preuves existent et que des processus valides d'évaluation des compétences et des savoirs professionnels sont réalisés dans les cours, que la validité curriculaire existe et qu'il y a une cohérence dans l'ensemble des évaluations d'un programme.

Construction et validation de modèles d'évaluation des compétences

Établir la *validité de construit curriculaire* nécessite de réaliser une exploration conjointe du référentiel de compétences, du cursus étudié et des dossiers universitaires. Cette exploration permet de préciser la structuration théorique attendue des évaluations qui sont rapportées au dossier académique et de soulever plusieurs hypothèses complémentaires à l'hypothèse H_1 . Il n'y a effectivement pas un seul et unique modèle qui peut rendre compte de la réalité et celui qui décrit le mieux l'ensemble des évaluations effectuées dans le cadre du cursus sera retenu. Le processus de sélection de modèles sera alors explicite. Le développement des différents modèles se fait à partir de l'analyse factorielle confirmatoire (AFC) (Brown, 2006). L'analyse factorielle est largement utilisée pour déterminer la validité de construit (par ex. : Bégin & Guay, 2010; Castillo, Leonardo, Cheniaux, & Landeira-Fernandez, 2010; De Ketele & Gérard, 2005; Issaieva & Crahay, 2010) et pour mettre en exergue certains effets liés à la méthode de mesure (par ex. : Chen, Rendina-Gobioff, & Dedrick, 2010). L'avantage de l'AFC sur l'analyse factorielle exploratoire est qu'elle permet d'évaluer la structuration des construits à partir des modèles élaborés *a priori*. Appliquée à l'analyse des dossiers universitaires, l'AFC sera l'outil privilégié dans le présent article pour statuer sur la *validité de construit curriculaire*.

Construction d'un modèle simple

Au départ, le référentiel de compétences lié à la formation des enseignantes et des enseignants (MEQ, 2001) demeure vague quant à la structuration théorique des compétences. À sa lecture, deux scénarios de structuration peuvent néanmoins être envisagés (voir figure 1). Ainsi, l'objectif ultime visé par le référentiel est le développement d'un professionnel compétent et cultivé (voir figure 1a). Cette orientation fondamentale se traduit d'abord en un modèle qui assume qu'il s'agit d'un construit latent unidimensionnel. Le construit est dit *latent* car la compétence est, par définition, non directement observable. Il est unidimensionnel car la présence d'un seul facteur est recherchée. Chaque cours devient alors un indicateur de la présence plus ou

moins forte de ce construit latent. Le développement mathématique de ce modèle s'inspire de celui présenté par Brown (2006). Il s'applique à la situation où une seule compétence est présente.

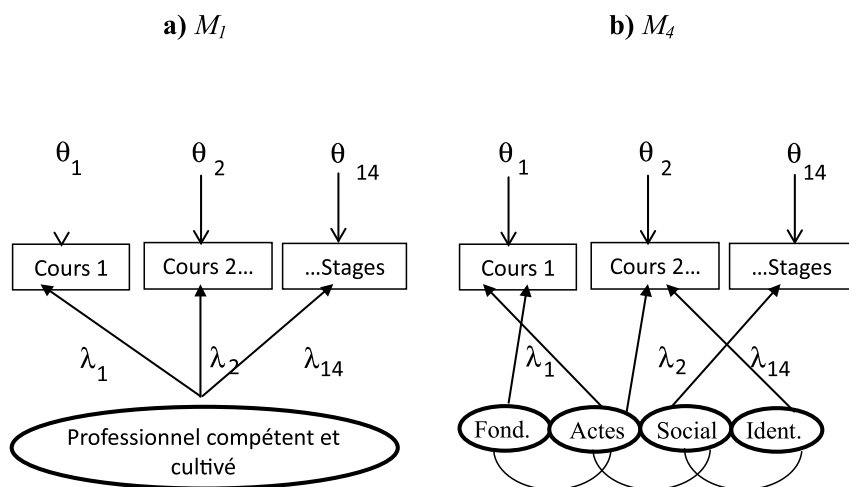


Figure 1. *Modèles M1 et M4*

Pour chacun des modèles, Θ représente la variance unique à chaque cours, λ est la saturation qui unit chaque facteur à chaque cours, Fond.= Fondements de l'éducation, Actes = Actes d'enseigner, Social = Aspects sociaux, Ident. = Identité. Une flèche identifie une saturation. Un carré identifie un cours et l'ovale un facteur latent. Un arc signifie que les unicités sont corrélées ; ainsi, les modèles M1 et M4 considèrent que les unicités sont indépendantes. Les points de suspension indiquent que d'autres cours s'ajoutent à ceux représentés.

Ainsi, la note académique (y) attribuée à un cours est une fonction du construit latent (η) et de la composante propre à chaque indicateur (ϵ). Cette composante propre est aussi parfois interprétée comme étant l'erreur de mesure.

L'équation (1) exprime cette fonction,

$$y = f(\eta, \epsilon) \quad (1)$$

où y est la note attribuée à un cours, η , le construit latent et ϵ , la composante spécifique à chaque indicateur.

La variance de la note, quant à elle, est également une fonction de la saturation (λ) qui unit la variable y au construit, la variance (ψ) du construit latent η , et la variance unique (Θ_ϵ) associée à l'indicateur. L'équation (2) exprime cette relation.

$$\text{Var}(y) = \lambda_y^2 \psi + \Theta_e \quad (2)$$

L'intérêt du modèle provient du fait qu'il opérationnalise la covariation qui existe entre les évaluations effectuées lors de cours différents. Celle-ci dépend de la saturation (λ) qui unit chacun des cours au construit latent et de la variance (ψ) de ce dernier. L'équation (3) exprime bien cette covariation (Cov),

$$\text{Cov}(y_1, y_2) = \sigma_{21} = \lambda_{y1} \psi \lambda_{y2} \quad (3)$$

où λ_{y1} et λ_{y2} représentent respectivement les saturations entre les variables y_1 et y_2 et le construit latent et ψ représente sa variance. Cette modélisation de la covariation entre deux cours amène l'énonciation de deux hypothèses fortes (H_{1a} et H_{1b}) liées à notre hypothèse principale.

H_{1a} Si l'ensemble des cours mesure un même construit latent associé à la culture et à la compétence d'un professionnel, certaines relations entre ces cours doivent exister.

Il sera possible de vérifier cette hypothèse à partir de la matrice de corrélation entre les cours du programme. De plus,

H_{1b} Si l'ensemble des cours mesure un même construit latent associé à la culture et à la compétence d'un professionnel, un modèle unidimensionnel constitué d'une seule variable latente pourra expliquer l'ensemble des relations entre les indicateurs (cours).

L'intérêt de ce modèle est d'ailleurs sa simplicité. Il permet d'expliquer la variance et la covariance entre les évaluations faites dans le cadre des différents cours d'un programme par la présence du construit latent et la composante spécifique à chaque cours.

Cette opération permet de répondre à une question d'intérêt : les différents cours concourent-ils à la mesure d'un professionnel compétent et cultivé ou divergent-ils ? Ainsi, bien que très simple, ce modèle unidimensionnel pose un jalon fondamental qui servira de base de comparaison face à des modèles plus complexes.

Construction d'un modèle à quatre composantes. Le référentiel précise que les programmes doivent évaluer douze compétences professionnelles et que celles-ci sont « interdépendantes » (MEQ, 2001, p. 57). Sur le plan psychométrique, cette prescription commande un modèle composé de douze dimensions interreliées. Cependant, un tel modèle nécessite un grand nombre de sujets et de cours pour permettre d'estimer ses paramètres constitutifs. Afin

d'éviter ce problème d'indétermination, un modèle intermédiaire, à quatre composantes, sera développé. Cette réduction du modèle de douze à quatre dimensions s'accompagne cependant d'une certaine perte d'information.

Ainsi, le référentiel suggère que les compétences se regroupent en quatre dimensions liées aux fondements, à l'acte d'enseigner, au développement de l'identité professionnelle et aux aspects sociaux de la profession. Il est donc possible qu'un modèle à quatre dimensions soit observé à partir des évaluations faites dans le cadre des activités. Ce modèle sera appelé M_4 (figure 1).

Dans ce modèle, le lien qui unit les cours aux compétences professionnelles est complexe et n'est pas le fruit du hasard. Comme il a été mentionné précédemment, une matrice *Cours-Compétence* condense les principaux éléments du programme et synthétise les objectifs de formation à atteindre par chacun des cours. La matrice *Cours-Compétences* utilisée est une adaptation de celle publiée par Harvey et Laroui (2006). La matrice originale a été réduite afin d'opérationnaliser ce modèle à quatre dimensions. La réduction en est une de logique. Elle permet d'établir la correspondance entre les cours et les quatre groupes de compétences définis par le référentiel (MEQ, 2001)⁴. Le tableau 1 présente la matrice qui sera utilisée. Celle-ci s'interprète comme suit. Le programme de formation prévoit que le cours *Didactique 1* mesurera la mobilisation de certains éléments (savoirs, ressources, compétences) des groupes de compétences associés aux *Fondements de l'éducation*, à *l'Acte d'enseigner*, et aux *Aspects sociaux*. Par contre, il n'évaluera pas celles impliquées dans le *Développement identitaire*. À ce stade-ci, les saturations associées à ce cours (λ_1 , λ_2 et λ_3) sont inconnues. L'objectif poursuivi par le présent article est de les estimer. Le développement mathématique de ce modèle dépasse cependant l'objectif du présent article.

L'hypothèse qui émerge de ce modèle est la suivante :

H_{1c} Si l'ensemble des cours mesure bien les quatre groupes de compétences définis par le référentiel, un modèle à quatre variables latentes pourra expliquer l'ensemble des relations entre les indicateurs (cours).

Il existe cependant une limite importante à la vérification de la validité de construit d'un programme à travers cette matrice. Celle-ci est une synthèse forcément incomplète du cursus. Ainsi, cette matrice ne détermine pas précisément ce qui est évalué dans le cadre de l'un ou de l'autre des cours en lien avec une compétence donnée ; alors qu'un cours peut évaluer quelques-unes des ressources qui seront ultérieurement mobilisées par une compétence,

un autre peut évaluer directement la compétence telle que mobilisée dans un contexte authentique. La matrice ne permet pas d'effectuer une telle distinction; *a priori*, elle considérera comme équivalente la contribution de l'un et de l'autre au développement de la compétence. *A posteriori*, des saturations plus faibles et des variances spécifiques fortes pourront être observées et, au final, si le phénomène est important, l'adéquation générale du modèle sera compromise et celui-ci sera rejeté.

Tableau 1
Matrice théorique du programme analysé

COURS	Fond. C1-C2	Acte d'enseigner C3-C6	Aspects sociaux C7-C10	Identité C11-C12
Didactique I	λ_1	λ_2	λ_3	
Didactique II	λ_4	λ_5	λ_6	
Stage I	λ_{30}	λ_{31}	λ_{31}	λ_{32}
Gest. de la classe	λ_7	λ_8	λ_9	λ_{10}
Stage II	λ_{33}	λ_{34}	λ_{35}	λ_{36}
Stage III	λ_{37}	λ_{38}	λ_{39}	λ_{40}
Stage IV	λ_{41}	λ_{42}	λ_{43}	λ_{44}
Org. scolaire	λ_{11}		λ_{12}	λ_{13}
TICE	λ_{14}	λ_{15}	λ_{16}	λ_{17}
Éthique	λ_{18}			λ_{19}
EHDAA	λ_{20}	λ_{21}	λ_{22}	
Apprentissage	λ_{23}	λ_{24}		
Évaluation		λ_{25}	λ_{26}	λ_{27}
Adolescence	λ_{28}		λ_{29}	

Note. Le paramètre λ correspond à la saturation à estimer pour chacun des cours pour chaque groupe de compétences.

Les stages, un lieu d'ancrage

La formation pratique est considérée comme un lieu privilégié pour évaluer les divers éléments d'une compétence (MEQ, 2001 ; Portelance, 2007). Plusieurs activités/cours de formation pratique sont présentes dans les programmes de formation et totalisent un minimum de 700 heures de présence en milieu scolaire. Dans notre analyse, ces cours de stage sont considérés comme un repère, un lieu d'ancrage privilégié. À ce titre, les évaluations qui sont faites dans les différents cours de stage devraient être très cohérentes

entre elles car elles mesurent des ensembles très similaires de compétences. De plus, il devrait exister une certaine cohérence entre ces cours et les autres du programme, car ils contribuent à développer des ensembles de compétences partiellement similaires. L'auteur soulève une hypothèse particulière en lien avec ce statut privilégié de la formation pratique.

H_{1d} Il y a aura une cohérence entre les évaluations réalisées lors des cours de formation pratique et celles qui se déroulent en milieu institutionnel.

Cette cohérence est nécessaire. Elle témoigne de la relation entre les différentes évaluations réalisées dans le cadre des cours du programme. La documentation (CAPFE, 2010; Prégent et al., 2009) encourage la mise en place d'une approche programme qui favorise cette cohérence entre les cours d'un programme. L'absence de toute cohérence (hypothèse H_{1a}) serait un indicateur négatif de la qualité du programme. L'hypothèse H_{1d} va au-delà de l'hypothèse H_{1a} et suggère que cette cohérence se manifestera par la présence conjointe des cours de formation pratique et des cours institutionnels dans des facteurs communs.

Méthodologie

Corpus de données

Les analyses portent sur le dossier académique de 164 universitaires diplômés du programme d'enseignement secondaire d'une université du Québec. La période échantillonnée se situe entre 2002 et 2010 et correspond à la période couverte par la dernière version agréée du programme de formation. L'invocation de ces données aux fins de la recherche a été approuvée par le comité d'éthique de l'université, et ce, dans le respect de l'anonymat des différents acteurs impliqués. L'échantillon se compose de 113 femmes et 51 hommes. L'âge moyen à la fin des études est de 26 ans.

Variables manifestes

Les évaluations (notée «E...A+») de chaque personne pour chacun des cours considérés du programme furent obtenues à partir de la base de données du service de la recherche institutionnelle de l'université où la recherche a eu lieu. Ainsi, les évaluations de 14 cours pour chacun des étudiants ont été recueillies et constituent les 14 indicateurs qui seront utilisés pour inférer la présence des construits latents des différents modèles⁵.

Description du curriculum

Le programme analysé se compose de 120 crédits et prépare les enseignants à œuvrer dans l'un ou l'autre des profils suivants : mathématiques, français, univers social, développement personnel, et sciences et technologies (voir Harvey & Laroui, 2006). Le curriculum est constitué des volets psychopédagogiques, formation pratique et disciplinaire. Le volet disciplinaire est unique à chaque profil et est constitué de cours qui visent à assurer que les futurs enseignants et enseignantes possèdent la culture propre à la matière enseignée.

Le volet formation pratique est composé de quatre stages (*Stages I, II, III, et IV*) totalisant 21 crédits et 700 heures de présence en milieu scolaire. Le volet psychopédagogique regroupe 30 crédits portant sur l'apprentissage et la cognition (*Apprentissage*), l'organisation scolaire (*Organisation scolaire*), les cours de didactique I et II (*Didactique I, Didactique II*), l'adaptation scolaire et la connaissance des élèves handicapés et en difficulté d'apprentissage (*EHDAA*), la gestion de classe (*Gestion de la classe*), l'évaluation des compétences (*Évaluation*), les technologies de l'information en éducation (*TICE*), l'adolescence et la vie adulte (*Adolescence*) ainsi que l'éthique et la profession enseignante (*Éthique*). Seuls les 14 cours explicitement identifiés aux volets psychopédagogiques et de formation pratique seront retenus pour l'analyse. Ce choix est justifié par le fait qu'ils assurent le développement de l'ensemble des compétences professionnelles. Les cours du volet disciplinaire, qui visent le développement de la *Compétence 1 Culture*, ont été exclus de la démarche, car les cours sont différents d'un profil disciplinaire à l'autre. Les cours optionnels du curriculum, qui ne sont offerts que sporadiquement, ont également été exclus.

Groupes de compétences évaluées

Les quatre groupes de compétences évaluées sont décrits dans le référentiel du ministère de l'Éducation du Québec (MEQ, 2001). Les compétences sont numérotées C₁ à C₁₂ pour faciliter leur identification. Le groupe *Fondements de l'éducation* concerne la culture de l'éducation (C₁) et la communication orale et écrite (C₂). *L'Acte d'enseigner* vise la conception (C₃) et le pilotage des situations pédagogiques (C₄), l'évaluation des compétences (C₅), et la gestion du groupe-classe (C₆). Les *Aspects sociaux* mobilisent l'adaptation de l'enseignement (C₇), les technologies de l'information et de la communication (C₈), la coopération avec l'équipe-école (C₉) et la coopération avec l'équipe pédagogique (C₁₀). Finalement, le groupe *Identité* inclut le développement (C₁₁) et l'éthique (C₁₂) professionnels.

L'analyse des données

Tel que précisé précédemment des analyses factorielles confirmatoires seront réalisées afin d'évaluer les paramètres des différents modèles (M_1 et M_4). L'algorithme basé sur le maximum de vraisemblance sera utilisé à partir des données brutes. Les analyses sont réalisées à l'aide du progiciel SAS 9.2 et de la procédure *Proc Calis*. Une contrainte a été appliquée sur les modèles à quatre dimensions afin que les corrélations entre les variables latentes doivent être supérieures ou égales à zéro. Il serait difficile d'expliquer que les variables latentes se développent en opposition l'une de l'autre.

Processus de sélection de modèles

Il existe plusieurs indices d'adéquation qui permettent de sélectionner un modèle parmi un ensemble (Brown, 2006 ; Harvey, 2011 ; Hélie, 2006 ; McCoach & Black, 2008). Les indices d'adéquation sont généralement basés sur la vraisemblance d'un modèle. Les indices utilisés dans le présent article sont le Chi-carré (χ^2), l'erreur moyenne quadratique d'approximation (*Root Mean Square of Error of Approximation*, RMSEA) et l'indice général d'adéquation (*General Fit Index*, GFI). D'autres indices existent mais ceux retenus sont parmi les plus utilisés. Les valeurs critiques sont $RMSEA \leq 0,06$, $GFI \geq 0,90$, et le χ^2 ne doit pas être significatif. De plus, les paramètres d'un modèle doivent être significatifs et se situer dans les limites attendues. Un modèle qui possède des saturations qui se situent à l'extérieur des valeurs attendues doit être rejeté. Par parcimonie, le modèle le plus simple qui satisfait les indices d'adéquation sera retenu.

Résultats

Statistiques descriptives

Le tableau 2 présente les statistiques descriptives de chacun des cours. Dans une distribution normale, les coefficients de symétrie et d'aplatissement sont de 0. L'inspection des coefficients des variables à l'étude suggère que celles-ci ne peuvent pas être considérées comme étant normalement distribuées. Brown (2006) souligne que l'absence de normalité de la distribution des indicateurs affecte notamment les indices d'adéquation et en particulier le χ^2 . Dans un tel cas, il recommande d'utiliser un indice approprié, soit le χ^2 ajusté ; ce qui sera fait.

Tableau 2
Statistiques descriptives pour chacun des cours

<i>Variable</i>	<i>Moy.</i>	<i>Écart type</i>	<i>N</i>	<i>Symétrie</i>	<i>Applat.</i>
Didactique 1	3,42	0,64	159	-0,75	0,62
Didactique 2	3,48	0,75	159	-1,70	3,75
Stage 1	3,94	0,48	130	-1,24	0,70
Gest. Classe	3,72	0,47	159	-0,99	0,68
Stage 2	3,69	0,58	154	-2,21	10,19
Stage 3	3,81	0,46	162	-1,19	1,92
Stage 4	4,11	0,33	157	-2,43	7,79
Org. Scolaire	3,61	0,64	149	-0,72	0,25
TICE	3,85	0,46	147	-1,40	2,82
Éthique	3,39	0,73	159	-1,30	2,87
EHDAA	3,94	0,55	152	-2,71	9,30
Apprentissage	3,71	0,62	148	-1,17	1,09
Évaluation	3,33	0,91	157	-1,18	1,05
Adolescence	3,42	0,80	151	-1,30	2,30

Le tableau 3 présente les coefficients de corrélation qui existent entre les différents cours. Ce tableau contient 91 coefficients en excluant la diagonale. De ce nombre, 58 sont positifs et significatifs, ce qui représente 63,7% des coefficients. Trente-deux sont non significatifs donc non différents de zéro, ce qui représente 35,2% des coefficients. Finalement, un coefficient est négatif et significatif. Ainsi, seuls deux cours, soient *Apprentissage* et *Organisation scolaire*, sont négativement corrélés. La corrélation négative entre ces deux cours est une forme d'incohérence. Cette situation suggère que les compétences qui sont construites dans une activité sont dévaluées dans l'autre. Fort heureusement, cette incohérence localisée n'invalide pas l'ensemble des analyses car près de 99% des cours possèdent une corrélation positive ou sont non corrélées. Dans l'ensemble, il existe une forme de cohérence dans les évaluations. À titre indicatif, les évaluations des cours *Didactique I et II*, qui mobilisent les mêmes groupes de compétences, sont fortement corrélées ($r = 0,41$). Les cours *Stages* sont également fortement intercorrélés. À l'opposé, le cours *Apprentissage* n'est pas corrélé avec les cours de formation pratique *Stages*. Cette absence de corrélation ne doit pas être interprétée comme une incohérence; ces cours mobilisent des groupes différents de compétences.

Seules les analyses factorielles confirmatoires permettront d'expliquer plus en profondeur ces interrelations. Néanmoins, à titre indicatif, la valeur du coefficient alpha de Cronbach a été calculée et se situe à $\alpha = 0,69$. Une valeur se situant entre $\alpha = 0,70$ et $\alpha = 0,80$ est jugée acceptable (George & Mallery, 2003). Sur la base de ces éléments, l'hypothèse H_{1a} peut être acceptée. Il existe des relations positives et significatives entre les évaluations des cours.

Tableau 3
Matrice de corrélation entre chacun des cours du programme

Cours	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Didactique 1	1,00													
2. Didactique 2	0,41	1,00												
3. Stage I	0,39	0,37	1,00											
4. Gest. Classe	0,33	0,27	0,29	1,00										
5. Stage II	0,28	0,38	0,43	0,14	1,00									
6. Stage III	0,33	0,31	0,21	0,19	0,40	1,00								
7. Stage IV	0,23	0,21	0,22	0,12	0,41	0,29	1,00							
8. Org. Scolaire	0,22	0,06	0,09	0,27	0,03	0,09	-0,08	1,00						
9. Tice	0,32	0,25	0,33	0,27	0,16	0,22	0,26	0,15	1,00					
10. Éthique	0,31	0,21	0,27	0,14	0,33	0,30	0,16	0,09	0,28	1,00				
11. EHDA	0,17	0,14	0,17	0,06	0,19	0,16	0,09	0,05	0,09	0,30	1,00			
12. Apprent.	0,10	0,16	0,09	-0,02	0,08	0,10	0,05	-0,20	0,16	0,26	0,15	1,00		
13. Évaluation	0,38	0,28	0,13	0,23	0,13	0,21	0,19	0,40	0,16	0,06	0,15	-0,12	1,00	
14. Adolescence	0,49	0,25	0,41	0,33	0,37	0,33	0,19	0,52	0,34	0,34	0,25	-0,11	0,35	1,00

Note. Les éléments en gras sont significatifs à $p \leq 0,05$ ou mieux.

Modèle M_1

Le modèle M_1 détermine si l'ensemble des cours mesure la présence d'un construit unidimensionnel. Le tableau 4 présente les indices d'adéquation obtenus. Ce modèle doit être rejeté car les indices d'adéquation ne sont pas acceptables (GFI = 0,79, RMSEA = 0,11). Par conséquent, les paramètres constitutifs ne seront pas présentés en détail dans cet article. Mentionnons brièvement que la saturation des indicateurs se situe entre [0,35 et 0,71] pour 12 des 14 cours. Deux cours (*Apprentissage* et *EHDA*) ne saturent pas avec les autres et possèdent des coefficients inférieurs à 0,30. Le rejet du modèle permet d'infirmier l'hypothèse associée H_{1b} . Un modèle à un seul construit

latent ne permet pas d'expliquer l'ensemble des interrelations entre les cours ; la notion de *professionnel cultivé et compétent* ne se résume pas à la mesure d'un concept unidimensionnel, du moins, telle qu'inférée à partir des évaluations de ces 14 cours.

Tableau 4
Indices d'adéquation obtenus pour les modèles étudiés

<i>Modèle</i>	<i>Paramètres</i>	<i>GFI</i>	<i>RMSEA</i>	χ^2_{aj}	<i>dl</i>	<i>P</i>
M1	28	0,79	0,11	205,00	77	0,00
M4	57	0,92	0,07	56,60	48	0,18

Modèle M₄

Le modèle M_4 est acceptable sur la base des indices d'adéquation (voir tableau 4). Les différents paramètres de ce modèle sont donc présentés en détails. Dans l'ensemble, la matrice présentée par le tableau 1 est validée à 75 %, car 75 % des saturations attendues se sont avérées significatives (voir tableau 5). L'hypothèse H_{1c} est corroborée ; un modèle à quatre construits latents permet de décrire les interrelations entre les cours. Afin d'explorer cet aspect plus en profondeur, le tableau 5 présente les saturations obtenues entre les cours et les quatre groupes de compétences.

Notons que plusieurs cours atteignent leurs objectifs et qu'ils saturent dans les groupes attendus de compétences. Pour le groupe *Fondements de l'éducation*, les saturations pour les cours *Didactique I et II*, *Stage III*, *TICE*, *Éthique*, *EHDAA*, *Apprentissage* et *Adolescence* sont positives et significatives. Les saturations pour les cours *Stage II*, *Stage IV* et *Organisation scolaire* ne sont pas significatives.

Pour le groupe *Acte d'enseigner*, les saturations pour les cours *Didactique I et II*, *Stage I, II, III et IV*, *TICE*, *Apprentissage* et *Évaluation* sont positives et significatives. Les saturations pour les cours *Gestion de la classe* et *EHDAA* ne sont pas significatives.

Pour le groupe *Aspects sociaux*, les saturations pour les cours *Didactique I et II*, *Stage I, II, III, IV*, *Gestion de la classe*, *Organisation scolaire*, *TICE*, *Évaluation* et *Adolescence* sont positives et significatives. Seule la saturation pour le cours *EHDAA* n'est pas significative.

Tableau 5
Matrice observée pour le programme analysé

<i>COURS</i>	<i>Fond.</i> <i>C1-C2</i>	<i>Acte d'enseigner</i> <i>C3-C6</i>	<i>Aspects sociaux</i> <i>C7-C10</i>	<i>Identité</i> <i>C11-C12</i>
Didactique I	0,44	0,38	0,47	
Didactique II	0,39	0,45	0,25	
Stage I	0,42	0,23	0,32	0,22
Gest. de la classe	0,30	0,18	0,40	-0,10
Stage II	0,22	0,24	0,32	0,48
Stage III	0,29	0,33	0,29	0,17
Stage IV	-0,01	0,33	0,35	0,62
Org. scolaire	0,09		0,65	-0,38
TICE	0,43	0,23	0,28	0,11
Éthique	0,70			0,12
EHDAA	0,37	0,13	0,06	
Apprentissage	0,30	0,30		
Évaluation		0,56	0,61	-0,37
Adolescence	0,48		0,75	

Note. Les saturations en gras sont significatives à 0,05 ou mieux

Pour le groupe *Identité*, seules les saturations pour les cours *Stage I, II*, et *IV* sont positives et significatives. Les saturations pour les cours *Gestion de la classe, Stage III, TICE*, et *Éthique* ne sont pas significatives. Pour deux cours, *Organisation scolaire* et *Évaluation*, des saturations significatives et négatives sont observées. Des saturations négatives indiquent une incohérence entre ce qui est évalué dans ces cours et les cours *Stages* qui possèdent des saturations positives pour cette composante.

Dans l'ensemble, l'hypothèse H_{1d} est corroborée ; il y a cohérence dans les évaluations de trois groupes de compétences sur quatre (tous sauf *Identité*). Certaines inférences sur le développement des compétences peuvent être effectuées à partir de ces trois groupes de compétences.

Il est à propos de vérifier jusqu'à quel point chacun des cours contribue à la mesure des construits communs par rapport à la mesure d'éléments qui leurs sont uniques. Bien qu'il n'y ait pas d'hypothèse spécifique à ce sujet, le tableau 6 fournit les données nécessaires à cette vérification. Il présente pour

Tableau 6
*Erreur, variance de l'erreur, variance totale et R-carré
 pour chacun des cours du programme*

<i>Variable</i>	<i>Variance expliquée</i>	<i>Unicité (€)</i>	<i>Variance de l'unicité (Θ)</i>	<i>Variance totale</i>	<i>R-carré</i>
Didactique 1	0,26	0,66	0,21	0,47	0,56
Didactique 2	0,25	0,76	0,35	0,60	0,42
Stage 1	0,09	0,78	0,15	0,24	0,39
Gest. Classe	0,07	0,84	0,17	0,24	0,30
Stage 2	0,09	0,74	0,11	0,20	0,44
Stage 3	0,06	0,83	0,14	0,20	0,31
Stage 4	0,05	0,62	0,03	0,08	0,61
Org. Scolaire	0,25	0,65	0,19	0,44	0,58
TICE	0,08	0,82	0,16	0,24	0,32
Éthique	0,22	0,71	0,21	0,43	0,50
EHDAA	0,04	0,92	0,20	0,24	0,16
Apprentissage	0,07	0,91	0,35	0,42	0,18
Évaluation	0,78	0,42	0,17	0,95	0,82
Adolescence	0,48	0,49	0,15	0,63	0,75
Moyenne	0,19	0,73	0,19	0,38	0,45

Note. Le R-carré est le pourcentage de variance expliquée par le modèle (variance expliquée/ variance totale). La variance totale est la variance de l'indicateur considéré. La variance de l'unicité est aussi appelée variance de l'erreur qui est propre à l'indicateur ciblé.

chaque cours, la composante propre à chacun (€), sa variance unique (Θ), la variance totale et le pourcentage de variance (R-carré) qui s'explique par le modèle. En moyenne, le modèle explique 45 % de la variance des évaluations réalisées dans le cadre des cours. La variance dans les évaluations des cours *Didactique I*, *Organisation scolaire*, *Stage IV*, *Éthique*, *Évaluation*, et *Adolescence* dépasse les 50 % et cette proportion est considérée très bonne. Pour six cours, le pourcentage expliqué se situe entre 30 % et 50 %. Il s'agit des cours de *Didactique II*, *Stage I, II et III*, *Gestion de la classe*, et *TICE*. Pour les cours *EHDAA* et *Apprentissage*, le taux est faible et se situe en deçà de 20 %. Ces cours mesurent donc à plus de 80 % des éléments qui leurs sont propres. En considérant les pratiques habituelles en psychométrie, des indicateurs qui ont une contribution trop spécifique sont généralement retirés d'un test. Dans le

cas présent, l'analyse porte sur un curriculum et, dans cette perspective, il n'est pas envisageable de retirer ces cours. Par contre, l'analyse suggère qu'une réflexion sur l'apport de ces cours à la mesure des compétences prévues au programme est nécessaire.

Par ailleurs, les coefficients de corrélation entre les groupes de compétences sont nuls. Ceux-ci peuvent donc être considérés indépendants. Ces groupes constituent quatre dimensions indépendantes associées à la notion de *professionnel compétent et cultivé*.

Discussion

L'analyse a permis de confirmer un certain nombre d'hypothèses. Tout d'abord, il existe des interrelations positives entre les évaluations réalisées dans le cadre des cours du curriculum ciblé (H_{1a}). Ces interrelations témoignent de la cohérence dans les évaluations du programme. Elles émergent du postulat que la validité curriculaire existe et que les évaluations réalisées sont cohérentes avec le contenu des cours et du curriculum. Ces interrelations ne s'expliquent cependant pas par la présence d'un facteur commun unique (rejet de H_{1b}). Tel que mentionné précédemment, le concept d'un professionnel compétent et cultivé ne se présente pas tel un concept unidimensionnel ; les compétences professionnelles ne se construisent pas telle une compétence unique.

En fait, l'adéquation du modèle à quatre dimensions permet de confirmer l'hypothèse H_{1c} . Dans l'ensemble, le curriculum analysé mesure de façon cohérente la présence de quatre groupes de compétences. Ici, le concept de professionnel compétent et cultivé devient multidimensionnel. Ces groupes indépendants de compétences sont liés à l'*Acte d'enseigner*, aux *Fondements de l'éducation*, aux *Aspects sociaux* et à l'*Identité* (MEQ, 2001). Les évaluations réalisées dans le cadre de trois des quatre groupes de compétences attestent de la cohérence des cours de formation pratique et de celles qui se déroulent en milieu institutionnel ; l'hypothèse H_{1d} est ainsi corroborée. Seule la dimension *Identité* est sujette à des incohérences entre certains cours institutionnels et de formation pratique. Par conséquent, aucune inférence quant au développement de ce groupe de compétences ne devrait être faite à partir du dossier académique. Chevrier, Gohier, Anadon, et Godbout (2007) ont déjà rapporté des perceptions différentes des dispositifs nécessaires pour développer cette composante en fonction de leur rôle et responsabilité dans le processus de formation. De telles différences entre les acteurs peuvent expliquer les incohérences observées dans cette composante.

Néanmoins, la corroboration des hypothèses H_{Ia} , H_{Ic} et H_{Id} pose les jalons qui permettent d'accepter l'hypothèse principale (H_I) du présent article. Il existe des interrelations entre les évaluations réalisées dans le cadre du curriculum; celles-ci s'expliquent par la présence attendue de quatre groupes structurés de compétences. De plus, il existe une cohérence entre les évaluations réalisées dans le cadre des cours offerts en institution et celles proposées dans le milieu professionnel. Ces éléments permettent d'établir la *validité de construit curriculaire*; les cours de ce programme mesurent bien ce qui est attendu dans le plan de formation, *i.e.* que les évaluations figurant au dossier universitaire permettent de porter un jugement valide et d'attester du développement de trois des quatre groupes attendus de compétences professionnelles.

Il n'est cependant pas possible d'affirmer que le curriculum analysé mesure bien l'ensemble des 12 compétences professionnelles, qui est l'objectif ultime de la formation. Il s'agit d'une limite à la présente étude. En effet, la validation d'un modèle à 12 dimensions interreliées est une opération complexe; elle dépasse la quantité d'information disponible pour l'étude. Pour atteindre un tel objectif, il aurait fallu disposer d'un plus grand nombre de cours et d'un plus grand nombre de finissantes et de finissants. Le présent article pose cependant les jalons qui permettront d'atteindre cet objectif dans des travaux ultérieurs.

La question qui se pose maintenant est d'expliquer pourquoi certains cours du programme ne participent pas tel qu'attendu à la structuration en quatre groupes de compétences. À cet égard, l'opposition de certains cours quant à la composante *Identité* est un exemple de tels résultats inattendus.

Ainsi, la méthodologie permet de déterminer ce qui, dans les évaluations, est attribuable à une mesure qui présente un bon niveau de validité et ce qui est spécifique au contexte d'un cours. Or, les résultats ont montré que les évaluations effectuées dans le cadre du programme possèdent une grande variabilité contextuelle; plus de 65 % de la variance⁶ est attribuable à la spécificité des cours. Plusieurs facteurs expliquent cette grande variabilité. D'abord, les cours d'un programme de formation ne sont pas des éléments standards. Les conditions dans lesquelles ils se déroulent sont variables; les ressources mobilisées ne sont pas identiques; il existe des fluctuations importantes dans les ressources professorales qui offrent ces cours. Par ailleurs, comme il a été déjà mentionné, la matrice du programme de formation est une synthèse du curriculum, et cette matrice a, de surcroît, été réduite à quatre groupes de compétences à des fins opérationnelles. En ce sens, les analyses ne

peuvent pas refléter les variations fines qui peuvent exister dans les évaluations. Par exemple, si les compétences servent de structure à l'ensemble de la formation, dans certains cours institutionnels, les étudiantes et les étudiants universitaires sont amenés à élaborer des savoirs professionnels (Buisse, 2012). Ces variations entre les cours engendrent de la variabilité non expliquée par l'analyse factorielle confirmatoire.

De plus, ce qui est attribuable à la variabilité contextuelle dans un modèle simple peut être pris en compte dans un modèle plus complexe. À titre indicatif, les construits mesurés sont appelés à évoluer durant la période prévue au curriculum, soit quatre années. Un modèle simplifié ne peut prendre en compte cette dimension évolutive du construit; certaines incohérences peuvent alors s'expliquer par l'horizon temporel qui sépare les cours l'un de l'autre. Il est à espérer, qu'avec des recherches ultérieures, des modèles longitudinaux, plus élaborés, pourront explorer ces facettes inaccessibles dans la présente étude.

Par ailleurs, en psychométrie, un indicateur qui ne mesure pas ce qu'il est supposé mesurer est généralement rejeté. Dans le cadre d'un programme, une telle interprétation trop sommaire pourrait engendrer une marginalisation des ressources associées à l'offre de certains cours et même le retrait de ceux-ci du programme. Or, il est possible que la spécificité de ces cours soit souhaitable en ce qu'ils contribuent à la diversification contextuelle du programme. Il faut donc nuancer l'interprétation.

Conclusion

Avec le foisonnement des programmes par compétences, les pratiques d'évaluation subissent des transformations importantes et, dans l'ensemble d'un cursus, il devient nécessaire de valider les mécanismes qui attestent du développement des compétences (CAPFE, 2010). Dans cette optique, le présent article a porté un jugement sur la *validité de construit curriculaire* d'un programme de formation en enseignement secondaire. À partir de l'analyse des dossiers académiques de finissants et de finissantes universitaires du programme, la démarche permet d'attester du développement de trois des quatre groupes de compétences. Elle complète les analyses orientées vers l'analyse curriculaire (Desjardins & Dezutter, 2009; Gauthier & Mellouki, 2006), les analyses spécifiques de certaines compétences (par ex. : Chevrier et al., 2007), et les évaluations de programmes internes ou externes, réalisées à partir de l'opinion de certains acteurs (par ex. : BRERIUQ, 2000a, 2000b). Puisque

l'analyse utilise les dossiers académiques comme substrat, elle se distingue des approches qui justifient la certification sur la base d'un bilan de compétences (voir Buysse, 2012) ou d'épreuves externes au programme.

Actuellement, le curriculum analysé atteint 75 % de ces objectifs car 75 % des saturations attendues dans la matrice du programme ont été observées. Cette statistique constitue un premier seuil de référence. Des comparaisons avec des programmes d'autres secteurs et d'autres universités (Cardin & Couture, 2006; Minier, 2006) permettront de juger si ce seuil peut être dépassé. Avec la mise en place d'une approche-programme centrée sur une plus grande cohérence entre les cours d'un cursus (Portelance, 2007; Prégent et al., 2009), l'hypothèse qu'il le soit est plausible.

Considérant qu'en éducation, la certification est basée sur la réussite des crédits au programme, porter un jugement sur la *validité de construit curriculaire* contribue positivement à légitimer et à valoriser la formation et répond ainsi au besoin exprimé de reconnaissance professionnelle (Jorro, 2009).

À moyen terme, la *validité de construit curriculaire* pourrait devenir une exigence des méthodologies d'évaluation de programmes (Ryan & Cousins, 2009; Wholey, Hatry, & Newcomer, 2010). Cependant, plus de recherches seront nécessaires afin d'établir de nouvelles normes et de nouveaux modèles qui permettront de baliser son interprétation. Dans un programme de formation, les construits mesurés sont multidimensionnels et évolutifs; les évaluations sont réalisées dans des conditions où il existe une grande variabilité contextuelle; les analyses nécessitent une quantité considérable de dossiers académiques; et enfin, les résultats obtenus sont dépendants de la matrice de formation analysée, ce qui signifie que l'ajout ou le retrait d'un ou de plusieurs cours pourrait donner des résultats différents. De plus, les normes actuelles de la psychométrie amènent à rejeter des indicateurs dont les saturations sont faibles. Comme il a été mentionné précédemment, les jugements basés sur les qualités psychométriques devront prendre en compte ces particularités inhérentes à l'étude de cette forme de validité. Dans la présente analyse, les cours avec des saturations faibles ou qui possèdent une grande variabilité contextuelle ne sont pas exclus du programme de formation; cependant, cette analyse peut susciter chez les personnes qui offrent ces cours une réflexion sur les dispositifs d'évaluation des compétences qu'ils utilisent.

NOTES

1. La forme masculine est employée afin d'alléger le texte.
2. La notion de cours fait référence à l'ensemble des activités créditées offertes dans le cadre du programme de baccalauréat. Ce terme inclut tant les cours institutionnels que les cours de stage se déroulant en milieu professionnel. Durant ces cours, des activités d'enseignement et d'évaluation ont généralement lieu.
3. À l'inverse, des résultats négatifs pourront s'expliquer soit par l'absence d'adéquation entre ce qui est évalué et ce qui devrait l'être ou le fait que le dossier universitaire n'est pas un instrument présentant un bon niveau de validité afin d'attester de l'atteinte des objectifs.
4. Ainsi, si un cours a contribué à développer l'une ou l'autre des compétences intrinsèques à un groupe, une indication positive est insérée dans la matrice. À l'opposé, si le cours n'a pas contribué à développer l'une ou l'autre des compétences, une absence de contribution apparaît.
5. Les notes littérales [A+...E] utilisées pour chacun des cours furent converties en cotes numériques en utilisant le barème suivant : « A+ \rightarrow 4,3 », « A \rightarrow 4,0 », « A- \rightarrow 3,7 » et ainsi de suite jusqu'à « E \rightarrow 0 ». L'échelle obtenue compte 12 niveaux et est généralement considérée et traitée comme une échelle d'intervalle. En effet, cette échelle possède un zéro absolu et les intervalles entre les niveaux sont considérés comme étant égaux. Par ailleurs, les dossiers universitaires rapportent une statistique telle que la moyenne de l'étudiant, une statistique réservée aux échelles d'intervalle. Les analyses effectuées dans le présent article respectent cette interprétation et traitent ces données comme provenant d'une échelle d'intervalle.
6. Ce résultat s'obtient en soustrayant $1-R^2$ moyen présenté au tableau 6.

RÉFÉRENCES

- American Educational Research Association (AERA) (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Bégin, C., & Guay, F. (2010). Validation de l'échelle de motivation à l'égard de l'orthographe (EMO). *Mesure et évaluation en éducation*, 33(1), 63-78.
- Bodin, A. (2006). Un point de vue sur PISA. *Gazette des mathématiciens*, 108, 54-59.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.
- Bureau du recensement étudiant et de la recherche institutionnelle de l'Université du Québec (BRERIUQ) (2000a). *Enquête auprès des premiers diplômés et diplômées du baccalauréat en enseignement secondaire*. Québec, QC : Université du Québec.
- Bureau du recensement étudiant et de la recherche institutionnelle de l'Université du Québec (BRERIUQ) (2000b). *Enquête auprès des employeurs des premiers diplômés et diplômées du baccalauréat en enseignement secondaire*. Québec, QC : Université du Québec.

- Buysse, A. A. J. (2012). Portfolio de développement professionnel et certification : Frictions constructives ou destructives. *Le bulletin de l'ADMEE-Europe*, 2012(1), 25-39.
- Cardin, J. F., & Couture, C. (2006). La refonte du programme du baccalauréat en éducation préscolaire et en enseignement primaire. In C. Gauthier & M. Mellouki (Eds.), *La formation des enseignants au Québec à la croisée des chemins* (pp. 121-160). Québec, QC : Les Presses de l'Université Laval.
- Castillo, C., Leonardo, M., Cheniaux, E., & Landeira-Fernandez, J. (2010). Psychometric properties and latent structure of the portuguese version of the Penn State worry questionnaire. *The Spanish Journal of Psychology*, 13(1), 431-443.
- Chen, Y.-H., Rendina-Gobioff, G., & Dedrick, R. F. (2010). Factorial invariance of a chinese self-esteem scale for third and sixth grade students: Evaluating method effects associated with positively and negatively worded items. *The International Journal of Educational and Psychological Assessment*, 6(1), 21-35.
- Chevrier, J., Gohier, C., Anadon, M., & Godbout, S. (2007). Construction de l'identité professionnelle des futures enseignantes : Dispositifs de formation présents et souhaités selon les acteurs responsables de la formation des maîtres au préscolaire et au primaire. In C. Gohier (Ed.), *Identités professionnelles d'acteurs de l'enseignement : Regards croisés* (pp. 137-168). Québec, QC : Presses de l'Université du Québec.
- Comité d'agrément des programmes de formation à l'enseignement (CAPFE)(2010). *La visite de suivi de l'agrément d'un programme de formation à l'enseignement : Cadre de référence et modalités d'application pour les visites de suivi de l'agrément qui seront effectuées entre 2010 et 2013*. Gouvernement du Québec, QC : Ministère de l'Éducation, du Loisir et du Sport.
- 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.
- Desjardins, J., & Dezetter, O. (2009). Développer des compétences professionnelles en formation initiale à l'enseignement préscolaire et primaire : regard sur l'organisation des programmes en contexte québécois. *Canadian Journal of Education*, 32(4), 873-902.
- Gauthier, C., & Mellouki, M. (2006). *La formation des enseignants au Québec à la croisée des chemins*. Québec, QC : Les Presses de l'Université Laval.
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference, 11.0 update (4th ed.)*. Boston, MA: Allyn & Bacon.
- Harvey, L. (2011). Hidden Markov models and learning in authentic situations. *Tutorials in Quantitative Methods for Psychology*, 7(2), 32-41.
- Harvey, L. (2009). L'échafaudage lors de la supervision en milieu professionnel : Étude des modalités et un modèle. *Mesure et évaluation en éducation*, 32(1), 55-83.
- Harvey, L., & Barras, D. (2008). Transfert des compétences et construction d'un langage d'action en situation de compagnonnage professionnels en soins infirmiers. *Revue des sciences de l'éducation*, 34(3), 665-687.
- Harvey, L., & Laroui, R. (2006). L'élaboration des nouveaux programmes de formation à l'enseignement secondaire : Enjeux et défis. In C. Gauthier & M. Mellouki (Eds.), *La formation des enseignants au Québec à la croisée des chemins* (pp. 237-252). Québec, QC : Les Presses de l'Université Laval.
- Hélie, S. (2006). An introduction to model selection : Tools and algorithms. *Tutorials in Quantitative Methods for Psychology*, 2(1), 1-10.

- Issaieva, É., & Crahay, M. (2010). Conceptions de l'évaluation scolaire des élèves et des enseignants : Validation d'échelles et étude de leurs relations. *Mesure et évaluation en éducation*, 33(1), 31-61.
- Jonnaert, P., Barrette, J., Boufrah, S., & Masciotra, D. (2004). Contribution critique au développement des programmes d'études : compétences, constructivisme et interdisciplinarité. *Revue des sciences de l'éducation*, 30, 667-696.
- Jorro, A. (2009). *La reconnaissance professionnelle : évaluer, valoriser, légitimer*. Ottawa, ON : Presses de l'Université d'Ottawa.
- Louis, R., Jutras, F., & Hensler, H. (1996). Des objectifs aux compétences, l'évaluation de la formation initiale des maîtres. *Revue canadienne de l'éducation*, 21(4), 414-432.
- McCoach, D. B., & Black, A. C. (2008). Evaluation of model fit and adequacy. In A. A. O'Connell & D. B. McCoach (Eds.), *Multilevel modeling of educational data* (pp. 245-273). Greenwich: Information Age Publishing Inc.
- Minier, P. (2006). Nouveaux programmes de formation à l'enseignement secondaire. In C. Gauthier & M. Mellouki (Eds.), *La formation des enseignants au Québec à la croisée des chemins* (pp. 191-220). Québec, QC : Les Presses de l'Université Laval.
- Ministère de l'Éducation du Québec (MEQ) (2001). *La formation à l'enseignement : Les orientations, les compétences professionnelles*. Québec, QC : Gouvernement du Québec.
- Ministère de l'Éducation, du Loisir et du Sport (MELS) (2006). *Le stage probatoire des enseignants et des enseignantes du préscolaire, du primaire et du secondaire : Document d'information*. Gouvernement du Québec, QC : Ministère de l'Éducation, du Loisir et du Sport.
- Portelance, L. (2007). Une vision partagée de la formation à l'enseignement. In L. Bélair, D. Laveault, & C. Lebel (Eds.), *Les compétences professionnelles en enseignement et leur évaluation* (pp. 197-212). Ottawa, ON : Presses de l'Université d'Ottawa.
- Prégent, R., Bernard, H., & Kozanitis, A. (2009). *Enseigner à l'université dans une approche-programme : Un défi à relever*. Montréal, QC : Presses internationales Polytechnique.
- Ryan, K. E., & Cousins, J. B. (2009). Introduction. In K. Ryan & J. B. Cousins (Eds.), *The Sage international handbook of educational evaluation* (pp. ix-xvii). Los Angeles, CA : Sage.
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Bruxelles : De Boeck.
- Wholey, J. S., Hatry, H. P., & Newcomer, K. E. (2010). *Handbook of practical program evaluation (Third edition)*. San Francisco, CA : Jossey-Bass.

Date de réception : 7 mars 2012

Date de réception de la version finale : 15 octobre 2012

Date d'acceptation : 17 octobre 2012