

L'évaluation :
levier pour l'enseignement
et la formation

Les 25, 26 et 27 janvier 2017
AGROSUP - Dijon

29^e COLLOQUE INTERNATIONAL

de l'Association pour le Développement
des Méthodologies d'Évaluation en Éducation

Axe 2 - L'évaluation : levier pour la (ré)ingénierie des formations

<i>Concevoir une cartographie des programmes universitaires de formation en fonction de leur qualité perçue et de leur réalité organisationnelle : un outil de pilotage pour améliorer les dispositifs de formation</i> Jean-Marc Braibant	p. 1
<i>L'évaluation du Plan Régional de Formation du réseau de l'Agence pour l'Enseignement Français à l'Étranger au Liban : quelles modalités pour un dispositif efficace ?</i> Rana Challah	p. 2
<i>Comprendre les niveaux de complexité des situations évaluatives utilisées pour évaluer : un levier indispensable de l'ingénierie de formation</i> Chantal Roussel	p. 4.
<i>Développer ses compétences sociales : l'exemple de la formation dispensée par TalentCampus</i> Catherine Archieri, Marlène Manach, Marianne Elies, Melanie Souhait	p. 5
<i>Evaluation d'un plan d'études sous l'angle de la relation formation-emploi</i> Olivier Delevaux	p. 6
<i>Développement et implantation d'un programme par compétences en Kinésiologie : enjeux pédagogiques, évaluation et perspectives</i> Adrien Cantat, Francis Lalime, Philippe Corbeil, Geneviève Charland, Daniel Turpin	p. 8
<i>Evaluation d'une plateforme internet d'apprentissage à l'écriture-lecture par des enseignants suivant une formation à son utilisation</i> Stéphanie Boéchat-Heer, Esther González-Martínez	p. 10
<i>Recyclage de questions à choix multiples d'épreuves certificatives de physique en items de tests formatifs en ligne</i> Pierre-Xavier Marique, Jean-François Van De Poël, Maryse Hoebeke	p. 12
<i>Exploration du développement de la compétence à évaluer des enseignants par la passation d'un test de jugement</i> Jennifer Smith, Eric Dionne	p. 14

Concevoir une cartographie des programmes universitaires de formation en fonction de leur qualité perçue et de leur réalité organisationnelle : un outil de pilotage pour améliorer les dispositifs de formation

Jean-Marc Braibant

L'évaluation des programmes par les étudiants est une pratique répandue à l'UCL. Le Service en appui à la qualité des programmes académiques (QOPA) met à la disposition des responsables académiques un questionnaire qui permet de recueillir la perception des étudiants à propos de la qualité de leur formation universitaire. Onze dimensions sont prises en compte :

- Satisfaction globale de la formation
- Clarté des objectifs
- Cohérence du programme
- Qualité de l'enseignement et de l'encadrement
- Pertinence des évaluations et des examens
- Adéquation de la charge de travail
- Développement de compétences transversales
- pour les méthodes pédagogiques
- Qualité des ressources et de l'environnement
- Implication personnelle
- Liens avec la pratique et le monde professionnel

Chacune de ces dimensions est évaluée à partir des réponses données à plusieurs items (degrés d'accord selon une échelle de Lickert). Ces enquêtes visent à procurer aux responsables et aux enseignants les moyens de diagnostiquer les forces et les faiblesses de leur programme en vue de définir les lignes d'action concrètes pour les améliorer. Elles sont réalisées à la demande des facultés et elles concernent les étudiants qui sont en fin de cycle de formation (BAC3 pour le 1^{er} cycle et Master2 pour le 2^e cycle).

Un autre service de l'UCL, le Service d'études (SET) rassemble différents indicateurs objectifs caractérisant l'organisation des programmes comme par exemple le nombre de cours offerts aux étudiants, la taille moyenne des amphis, la charge horaire moyenne de l'étudiant, le pourcentage des activités pratiques, etc. Ces indicateurs sont nombreux et parfois redondants. Une analyse en composantes principales a permis d'en retenir 3 :

- le nombre de cours offerts qui reflète la diversité de l'offre de formation proposée aux étudiants,
- la taille moyenne des amphis, c'est-à-dire le nombre d'étudiants inscrits en moyenne dans un cours;
- la proportion des heures pratiques (TP, TD, labos, projets, ...) par rapport au nombre d'heures totales dispensées.

Dans le cadre d'une étude exploratoire, la collaboration des deux services a permis de croiser les deux banques de données et de répondre à la question suivante : existe-il un lien entre la qualité perçue des programmes de 1^{er} cycle par les étudiants et ces trois indicateurs objectifs ? Les résultats (Braibant & Tempels, 2015) ont montré que :

- la diversité de l'offre de formation n'est corrélée positivement avec aucune dimension liée à la perception de la qualité de la formation;
- la taille moyenne des amphis est corrélée négativement avec la qualité de l'encadrement et la pertinence des évaluations certificatives;
- le poids relatif des activités pratiques est corrélé de manière significative à la qualité perçue des programmes pour 7 dimensions sur 11. La corrélation avec la satisfaction globale est élevée : 0,65.

Ce dernier résultat confirme l'intérêt d'associer aux cours magistraux un nombre élevé d'activités

pratiques par petits groupes (Bédard & Viau, 2001, Johnson & Johnson, 1999). Cette étude suggère aussi qu'il existe des différences importantes entre les programmes en termes de qualité perçue et d'organisation pédagogique. Quelles sont les caractéristiques des programmes qui privilégient les activités pratiques et qui sont appréciés par les étudiants ? A contrario, quelles sont les caractéristiques des «autres» programmes, ceux qui accordent une place plus importante aux cours magistraux ? S'agit-il d'un choix délibéré ou s'agit-il d'un choix par défaut que l'on pourrait mettre en relation avec des effectifs plus importants d'étudiants et/ou des ressources limitées d'encadrement ?

Répondre à ces questions constitue l'objectif de la présente étude. Pour ce faire nous avons utilisé la méthode de WARD qui consiste à classer et à regrouper des sous-ensembles relativement homogènes au sein d'une population de départ caractérisée par des variables quantitatives. Dans cette étude, nous avons retenu 7 indicateurs pour caractériser les 29 programmes de 1^{er} cycle évalués par les étudiants ces 5 dernières années (n =1768; taux de réponse de 53,2 %) :

- qualité perçue de la formation par les étudiants
- nombre de cours offerts dans le programme
- taille moyenne des amphithéâtres
- nombre d'heures pratiques / nombres d'heures totales
- nombre d'étudiants par enseignant c'est-à-dire le nombre moyen d'étudiants auxquels un enseignant fait face en moyenne dans le programme
- taux d'échec en 1^{ère} année de bachelier
- nombre moyen d'années pour être diplômé

L'analyse via la méthode de Ward a permis de dégager quatre profils différents de programmes.

• Le cluster 1 regroupe des programmes dont la qualité est fortement appréciée par les étudiants de BAC 3. Confirmation des résultats de la 1^{ère} étude : ce sont des programmes qui privilégient les activités pratiques et l'offre de formation est peu diversifiée. On observe que la taille moyenne des amphithéâtres est basse et que le taux d'encadrement est très favorable. Notons enfin qu'il s'agit de programmes sélectifs : le taux d'échecs en BAC 1 est plus important en moyenne que dans les autres clusters et la durée de diplomation est longue

• Le cluster 2 regroupe des programmes pour lesquels les étudiants expriment également leur satisfaction. Le profil de ces programmes diffère de ceux du Cluster 1 sur plusieurs points : il s'agit de programmes peu sélectifs rassemblant un nombre élevé d'étudiants et un faible taux d'encadrement. Ici, on propose une grande diversité de cours alors que le nombre d'activités pratiques est peu élevé.

• Le cluster 3 rassemble les programmes dont la qualité perçue est la plus basse. Pourtant, les étudiants y sont peu nombreux, le taux d'encadrement est élevé et le taux d'échecs en 1^{ère} année est bas. On note que l'offre de formation est très diversifiée et que les activités pratiques prennent une place peu importante dans la formation.

• Le cluster 4 reprend également des programmes peu satisfaisants. Contrairement au cluster 2, ils rassemblent un nombre très élevé d'étudiants encadrés par un nombre réduit d'enseignants. L'offre de formation n'est pas très diversifiée même si les TP sont relativement nombreux. Ces programmes enregistrent un taux d'échec relativement bas en BAC1.

Cette étude permet de dessiner une cartographie des programmes de formation universitaire en fonction de leur qualité perçue et de leur réalité objective. Cet outil de pilotage montre que des marges de manœuvre existent en termes de choix organisationnels et pédagogiques pour optimiser les dispositifs de formation

L'évaluation du Plan Régional de Formation du réseau de l'Agence pour l'Enseignement Français à l'Étranger au Liban : quelles modalités pour un dispositif efficace ?

Rana Challah

L'évaluation des dispositifs de formation constitue un enjeu stratégique et une valeur ajoutée pour l'institution dans la mesure où elle permet aux décideurs de la formation de l'adapter aux besoins des

usagers et de mieux gérer les budgets ou l'investissement de la formation (Barbier, 1985; Barbier, J.-M., Chaix, M.-L., Demailly, L., 1994; Sainsaulieu, 1999; Dennery, 1997, 2001). En ce sens, l'évaluation de la formation peut aussi contribuer à l'amélioration de l'offre de la formation, à la certification de l'atteinte des objectifs visés et, par conséquent, à l'amélioration de la conception de la formation. En vue de vérifier la cohérence des actions de formation avec les objectifs énoncés dans le vade-mecum de la formation continue de son service pédagogique, l'Agence pour l'Enseignement Français à l'Étranger évalue, par le biais d'un questionnaire, l'impact du Plan Régional de Formation (PRF). Le PRF regroupe un catalogue annuel de l'AEFE, proposant des stages qui peuvent avoir lieu dans toutes les zones du réseau mondial, ainsi que des stages spécifiques à chaque zone. Nous nous intéressons particulièrement au réseau de l'AEFE au Liban. En effet, selon le vade-mecum du service pédagogique de l'AEFE (2013), l'évaluation de la formation continue doit rendre compte de «la traçabilité du parcours de formation des enseignants [et de] l'évolution des pratiques professionnelles» (AEFE, 2013, p. 2). L'objectif de notre communication est de questionner les modalités d'évaluation du PRF au sein du réseau de l'AEFE au Liban et l'impact de ce dispositif sur les pratiques professionnelles des enseignants. Dans le cadre de notre recherche doctorale, nous avons élaboré un questionnaire à échelles de type Likert que nous avons adressé aux coordonnateurs de discipline et aux enseignants du réseau afin de recueillir leurs représentations quant aux modalités de mise en œuvre de la formation et à l'impact de ce dispositif de formation sur leurs pratiques professionnelles. Plus précisément, nous avons effectué un traitement statistique de 423 questionnaires remplis par des enseignants et coordonnateurs de discipline qui travaillent au sein du réseau. Nous avons également analysé le contenu du formulaire d'évaluation de l'AEFE, afin de faire ressortir les éléments évalués, pour vérifier si ces derniers correspondaient bien aux objectifs annoncés par l'AEFE. Plus précisément, nous avons analysé 27 fiches de synthèse remplies par 483 enseignants.

Si l'enjeu de l'évaluation est de garantir la qualité de la formation et de vérifier l'atteinte des objectifs visés par les décideurs de formation qui assurent une régulation de la formation afin qu'elle corresponde aux orientations stratégiques de l'institution et aux attentes des usagers, les résultats de l'enquête reflètent l'existence d'une contradiction dans la manière dont les enseignants évaluent la formation continue. En effet, les résultats de notre enquête par questionnaire reflètent une appréciation peu positive voire négative de ce dispositif de formation continue alors que l'évaluation des stages confirme une satisfaction générale de celui-ci. L'évaluation proposée par l'AEFE est-elle considérée comme étant une formalité aux yeux des enseignants? Est-elle fiable? Par ailleurs, l'examen des résultats nous permet de constater que les modalités ainsi que les contenus de l'évaluation ne correspondent pas aux objectifs annoncés par l'AEFE. Il semble que cette évaluation est faite à des fins statistiques au détriment d'un réel suivi des acquis des enseignants. Comment penser l'évaluation du PRF? Quel(s) modèle(s) d'évaluation pour garantir la qualité et l'efficacité de ce dispositif de formation? Nous proposons un modèle d'évaluation de la formation dans lequel la conception ainsi que l'évaluation de la formation se présentent comme deux processus imbriqués.

Bibliographie

AEFE. Vade-mecum du SP (Service Pédagogique) : RÉUSSIR relatif à « l'élaboration, l'organisation et la réalisation du plan régional de formation. Récupéré le 5 juillet 2014 de <http://www.aefe.fr/sites/default/files/asset/file/2013-09-vade-mecum-reussir-un-prf>.

Barbier, J.-M. (1985). L'évaluation en formation. Éducation et formation. Paris : PUF.

Barbier, J.-M., Chaix, M.-L., Demailly, L. (1994). Éditorial. Recherche et Formation, 17, 5-8.

Dennery, M. (1997). Organiser le suivi de la formation : méthodes et outils. Paris : ESF.

Dennery, M. (2001). Évaluer la formation : des outils pour optimiser l'investissement formation. Paris : ESF.

Mots clés : évaluation, dispositif du PRF, AEFE

Comprendre les niveaux de complexité des situations évaluatives utilisées pour évaluer : un levier indispensable de l'ingénierie de formation

Chantal Roussel

Actuellement, la cohérence des formations professionnelles demeure au centre des préoccupations pédagogiques des équipes chargées de les concevoir ou de les dispenser, car elle se fonde sur l'organisation des compétences à développer. En lien avec la structure des programmes, l'évaluation tient une place centrale pour rendre compte des niveaux de développement atteints par les apprenants à différentes étapes de leur parcours de formation.

En examinant les enjeux présents dans diverses études recensées, nous observons un portrait disparate, marqué par la diversité des angles et des points de vue. Dans ces « compositions », il ressort une chose qui demeure le plus souvent implicite, c'est-à-dire celle à laquelle nous nous intéressons dans le cadre de nos recherches. En réalité, on semble très souvent occulter ce qui compose les tâches d'évaluation. Ce constat est surprenant, en particulier lorsque l'on comprend que ces tâches constituent les attentes exigées envers les élèves ou les étudiants, au terme d'une séquence d'apprentissages ou en fin de formation. Cette conclusion nous permet de penser qu'il existe des difficultés potentielles ou bien réelles sur le plan des pratiques d'évaluation, puisque celles-ci sont fortement guidées par la composition des situations évaluatives présentées.

La recherche que nous avons menée durant notre doctorat nous a conduite à examiner les niveaux de complexité des situations évaluatives (SÉ) de la première année d'un programme d'études de la formation professionnelle au secondaire. Par la suite, une fois la thèse soutenue, nous avons poursuivi cette recherche avec l'analyse des situations évaluatives de la seconde année de ce programme. Lors de ces recherches, nous avons pris appui sur des éléments théoriques présentés par Roegiers (2007) et permettant de baliser l'interprétation des situations évaluatives analysées. Pour résumer le propos de ces auteurs, une compétence professionnelle intègrerait l'ensemble des ressources situées à des niveaux inférieurs : des savoir-faire professionnels. Ces savoir-faire professionnels combinent à leur tour un ensemble de savoir-faire techniques et fonctionnels, eux-mêmes s'organisant à l'aide de connaissances, de capacités techniques et fonctionnelles, ainsi que de savoirs d'expérience. Selon Roegiers *et al.* (2010), « [T]out comme pour les savoir-faire professionnels, les compétences professionnelles ne peuvent s'acquérir qu'en situation, selon une intégration progressive [...] » (p. 23). La compétence se maîtrise donc seulement en contexte réel de pratique.

Par ailleurs, pour étudier les situations évaluatives, nous avons conçu des outils d'analyse en nous appuyant sur différents écrits théoriques et scientifiques discutant de l'évaluation de performance et de l'évaluation authentique. Ces outils prennent la forme de trois différents continuums. Ils ciblent des éléments précis des situations évaluatives : les ressources internes et externes mises en œuvre, l'articulation qui existe entre ces ressources, puis le contexte employé au moment de leur mise en œuvre. Enfin, ces continuums sont assemblés pour dévoiler des configurations particulières des situations évaluatives examinées.

Pour comprendre les niveaux de complexité des situations évaluatives des deux années du programme SASI, correspondant à 31 compétences déclinées dans la formation, une stratégie méthodologique recourant à l'étude de cas instrumentale a été déployée. Dans cette séquence méthodologique, nous avons eu recours à des analyses documentaires, à des entretiens individuels et de groupe, ainsi qu'à la tenue d'un journal de bord par la chercheuse. Ces trois modalités se justifient parce que nous souhaitons porter un regard que l'on pourrait caractériser de « triangulé » sur les ingrédients composant l'objet : situations évaluatives de compétences. Dans ce contexte de recherche, des analyses successives ont été menées, afin de parvenir à questionner plus finement les enseignants.

Lors de notre recherche doctorale, les niveaux de complexité des situations évaluatives de la première année du programme SASI touchaient *l'exploration, l'appropriation, l'application procédurale et l'intégration partielle*. À la suite de cette recherche doctorale, nous avons poursuivi l'examen des

situations évaluatives de la seconde année du programme SASI, afin d'en analyser les autres compétences et d'interpréter leurs niveaux de complexité. On y retrouve des niveaux similaires à ceux de la première année, dont l'évaluation de l'appropriation utilisant des examens de connaissances théoriques et de l'application procédurale employant des évaluations de techniques de soins en contexte de laboratoire. En rapport avec le niveau d'intégration partielle, il se matérialise par des situations évaluatives utilisées en contexte simulé (en laboratoire et réel, mais pas en contexte de pratique) et en contexte réel d'exercice professionnel.

Concernant le niveau d'intégration complète des apprentissages, les résultats de cette recherche montrent qu'il ne semble pas entièrement utilisé. Les élèves se retrouvent bel et bien en contexte réel de soins, mais les situations déployées pour l'évaluation ne revêtent pas exactement les mêmes niveaux d'exigence, c'est-à-dire qu'ils ne mobiliseraient pas exactement les mêmes ressources que celles impliquées sur le marché du travail (en terme de capacité de réponse à un nombre élevé de patients, de rapidité d'exécution liée à certaines tâches de la profession d'infirmière auxiliaire, etc.). Des questions visant à éclairer ce constat ont alors été posées aux enseignants chargés de cette formation pour en comprendre les subtilités. Leurs réponses provoquent des interrogations et des discussions, car ces derniers expliquent qu'un écart demeure entre le niveau atteint en fin de formation et celui exigé par le milieu professionnel. Pour eux, plusieurs facteurs contextuels expliquent cette situation. Entre autres le type de région où ils se trouvent, la capacité et la disponibilité d'accueil des milieux de ces stagiaires, les ratios enseignant/élèves pour l'encadrement, ainsi que la structuration des évaluations en tant que telles.

Les résultats suggèrent que des pistes de recherche future doivent contribuer à la compréhension des niveaux de complexité des situations évaluatives, mais également à l'arrimage des situations évaluatives organisées les unes par rapport aux autres dans un plus vaste projet de développement des compétences. Cette considération nous semble incontournable dans des chantiers visant la production d'une ingénierie de formation. Cette recherche apporte ainsi des éclairages pertinents et novateurs au sujet de l'évaluation des compétences en formation professionnelle.

Mots clés : situations évaluatives, compétences, niveaux de complexité

Session 3 - Atelier 3.5

Jeudi 26 janvier 2017 – 08h30-10h00

Développer ses compétences sociales : l'exemple de la formation dispensée par TalentCampus¹

Catherine Archieri, Marlène Manach, Marianne Elies, Melanie Souhait

Depuis quelques années, les notions de compétences sociales et compétences comportementales sont des critères d'importance en matière de recrutement, d'orientation, voire de réorientation professionnelle. Ceci est lié à l'évolution des situations professionnelles, des métiers et à la nécessité de travailler de plus en plus en réseau (Bellier, 2000). Dans cette perspective, TalentCampus propose un nouveau modèle de formation fondé sur le développement des compétences sociales. Le dispositif de formation proposé s'adresse à un public varié (étudiants, salariés, demandeurs d'emploi, de 18 à 60 ans) avec la finalité de révéler leurs compétences relationnelles et collaboratives.

Dans un souci d'efficacité et d'innovation constante, TalentCampus cherche à évaluer de manière continue les différents modules qu'il propose. Pour cela, le dispositif met en place un suivi à court et à moyen termes des stagiaires à travers des questionnaires en ligne. Cependant, certains éléments autour de l'incidence de la formation ne sont pas renseignés par cette méthode d'évaluation. C'est pourquoi

¹ Ce travail a bénéficié d'une aide de l'Etat gérée par l'Agence Nationale de la Recherche au titre du programme investissements d'avenir portant la référence n°ANR-11-IDEFI-0035

l'observation de l'activité des stagiaires est envisagée comme une autre piste pour renseigner sur les processus de développement des stagiaires au cours des sessions de formation.

Le questionnement formulé par TalentCampus en direction du CREAD a donc été le suivant : quelles sont les transformations repérables chez le stagiaire, relatives aux compétences relationnelles, dans le cadre de la formation? Quelles sont conditions les plus favorables à ces transformations ?

Pour répondre à ces questions, l'équipe scientifique du CREAD a mis en place un observatoire de recherche fondé sur une approche «orientée activité», adapté du cadre théorique et méthodologique du «cours d'action» de Theureau (2004, 2006) et développé dans le courant de l'ergonomie des situations de formation (Guérin, 2012, Zeitler, 2011). L'analyse de la construction de l'expérience de deux stagiaires volontaires a été réalisée avec pour objectif de rendre compte des incidences de l'environnement de formation sur leur transformation-développement. En parallèle, pour documenter l'activité du formateur, des entretiens semi-directifs ont été menés et analysés en mobilisant l'approche sociotechnique développée par Albero (2010).

La finalité de cette enquête était de repérer les différents modes d'engagement des stagiaires, de les recouper avec ceux du formateur pour au final en évaluer la congruence.

Les résultats issus de l'analyse des entretiens menés auprès des stagiaires et du formateur, associés aux résultats issus de l'analyse des questionnaires en ligne, portent essentiellement sur l'intérêt des outils utilisés en formation, la dynamique de groupe et le cadre posé par le dispositif. Tout d'abord, cette enquête a montré que les outils mis à disposition dans les actions de formation constituaient des ressources utiles qui ont favorisé l'introspection, la mise en valeur de soi au sein du collectif et l'échange autour de valeurs communes, telles que la bienveillance. Ces outils, essentiellement en lien avec des compétences intra et interpersonnelles, ont également été jugés utiles parce qu'ils ont été considérés par les stagiaires comme signifiants et mobilisables dans la vie quotidienne. Ensuite, cette analyse a montré que la diversité des actions de formation autour de la mise en scène de soi et de l'intelligence collective permettait aux apprenants une prise de conscience. Le collectif a donc été considéré comme facilitateur, voire déclencheur d'une transformation chez les stagiaires. Enfin, cette analyse a montré que le dispositif de formation, à travers sa logique d'action, influençait l'engagement des stagiaires dans la formation et donc leur transformation-développement.

A l'issue de cette étude et dans une perspective d'évolution de la formation, des pistes de réingénierie pédagogique et de formation ont été proposées.

Références bibliographiques

Albero, B. (2010). La formation en tant que dispositif : du terme au concept. In Charlier, B., Henri, F. (dir.). La technologie de l'éducation: recherches, pratiques et perspectives. Paris : PUF, 47-59

Bellier, S. (2000). Compétences comportementales : appellation non contrôlée. In Bellier, S. Compétences en action, expérimentations, implications, réflexions pratiques. Editions Liaisons

Guérin, J. (2012). Activité collective et apprentissage : de l'ergonomie à l'écologie des situations de formation. Paris : L'Harmattan

Theureau, J. (2006). Le cours d'action : méthode développée. Toulouse : Octarès Editions

Zeitler, A. (2011). Les apprentissages interprétatifs: interprétation en action et construction de l'expérience. Paris : L'Harmattan

Mots clés : compétences sociales, pédagogie active, analyse de l'activité humaine

Evaluation d'un plan d'études sous l'angle de la relation formation-emploi

Olivier Delevaux

Tout en poursuivant une mission de transmission de certaines valeurs, dont le «vivre ensemble» n'est pas la moindre, le système éducatif doit s'adapter à l'évolution sociétale et à de nouvelles exigences. Il doit, par exemple, intégrer la demande accrue de transparence, adopter une orientation prenant en compte les

attentes des bénéficiaires et des différents partenaires, eux-mêmes de mieux en mieux informés et davantage parties-prenantes. Les parents aspirent légitimement à l'autodétermination et à l'appropriation dans le processus de scolarisation. La problématique de la place des élèves ayant des besoins éducatifs particuliers au sein de l'école publique représente un élément emblématique de l'évolution de la profession. Non seulement elle vient questionner de manière incontournable les rôles et identités professionnels des différents intervenants, mais elle contraint à des pratiques accrues de collaborations, que ce soit avec les parents ou entre les différents intervenants.

Dans ce contexte en mutation, et soumise à des impératifs nouveaux, les formations d'enseignants doivent se préoccuper de préparer les futurs diplômés à relever les défis de la profession. Elles ont donc évolué, visant des objectifs nouveaux. Certaines conditions semblent par ailleurs requises pour favoriser «des pratiques d'enseignants professionnels qui répondent à la demande du terrain, qui s'adaptent aux situations éducatives conjoncturelles rencontrées et qui évoluent vers la réussite du plus grand nombre d'élèves, c'est-à-dire vers une plus grande démocratisation de l'éducation» (Altet, 2002, p. 14). Il nous semble dès lors particulièrement intéressant de nous pencher sur l'évaluation de ces nouvelles formations et de chercher à établir si elles atteignent les buts qu'elles se sont fixées en termes d'adaptation à l'évolution du métier d'enseignant.

Concernant précisément les évaluations des formations à l'enseignement, Périsset et Lehmann (2010, p. 16) constatent que ce «volet de la recherche qu'imposent les nouvelles gouvernances, soit l'évaluation et l'auto-évaluation des dispositifs de formation n'est pas encore vraiment entré dans les usages.» Les auteurs font part de deux raisons d'étonnement. «La première, c'est que les HEP comme les Universités ont développé des départements dits d'«Assurance Qualité». Mais les enquêtes évaluatives menées dans ce contexte restent confidentielles, confinées à un usage administratif interne. [...] L'autre sujet d'étonnement, lié certes au premier, relève d'absence de prise en main, par la profession ou du moins par les formateurs, de l'évaluation des dispositifs de formation dans le but d'en reconnaître les effets (s'il y en a) et de les améliorer constamment, en connaissance de cause, de manière volontaire et non imposée par une hiérarchie ou par décision politique.» A la Haute école pédagogique du canton de Vaud, l'évaluation des formations prend essentiellement la forme d'une évaluation des enseignements par les étudiants et la disponibilité d'autres indicateurs, provenant notamment d'évaluations externes, est faible.

Ces constats nous ont encouragé à effectuer une évaluation du Master en enseignement spécialisé en nous intéressant prioritairement à la perception, par des diplômés récents ainsi que par des employeurs, d'une part de l'adéquation des compétences proposées dans le référentiel en regard des défis de la profession, d'autre part de la capacité de la formation à développer l'ensemble des compétences visées. Attentif à ne pas céder à une vision mécaniste de l'ingénierie de formation (Astier, 2007; Lambelet, 2009), ni à une perspective adéquationniste de la relation formation-emploi (Ourliac, 2014), nous avons pris en compte certains éléments contextuels déterminants. Il s'agit tout d'abord d'une formation en emploi faisant suite à une première formation à l'enseignement généraliste ou à une autre formation dans un domaine d'études considéré comme voisin. Les diplômés en enseignement spécialisé forment donc un groupe hétérogène si l'on considère leur formation initiale et leur parcours professionnel antérieur. Une autre caractéristique à prendre en compte est la vocation «généraliste» d'une formation destinant les enseignants spécialisés à intervenir tant en contexte intégratif qu'en contexte séparatif, et avec tout type d'élèves relevant de mesures de l'enseignement spécialisé.

En lien avec ces particularités, notre étude exploratoire s'est focalisée sur la relation formation-emploi en prenant en compte principalement deux des indicateurs de qualité d'une formation mis en évidence par Gérard (2001) : l'équilibre, considéré sous l'angle de la capacité de développer l'ensemble des compétences du référentiel, et l'efficacité externe «qui prend en compte les produits ou effets engendrés par le système de formation observés hors de ce système lui-même» (p.57), en particulier ici la capacité à travailler dans les différents contextes professionnels mentionnés et à contribuer à la mise en œuvre de la politique publique en matière de scolarité intégrative.

Les données recueillies montrent globalement l'adéquation des compétences retenues dans le référentiel en regard des défis professionnels, certaines compétences étant cependant jugées plus pertinentes que

d'autres. Par contre, si l'on considère la perception chez les diplômés et les employeurs de la capacité de la formation à développer ces compétences, on s'aperçoit que certaines lacunes sont mises en évidence. De même, les objectifs de formation fixés par la réglementation suisse sont jugés diversement atteints.

Le croisement de la perception des compétences les plus importantes en regard des défis de la profession avec la perception des compétences les plus développées par la formation laisse entrevoir une problématique plus importante, à savoir que certaines compétences, jugées très importantes pour le métier, seraient faiblement développées par la formation. Une autre surprise vient de l'importance constatée de la formation antérieure et du contexte d'emploi avant et durant la formation sur la perception des compétences développées durant la formation.

Le premier constat -à savoir les lacunes perçues de la formation- pourrait a priori déboucher assez simplement sur des solutions consistant en une réorientation de certains contenus et la mise sur pied de modules renforçant les dimensions jugées lacunaires. Le second constat quant à lui incite à s'interroger sur des dispositifs susceptibles d'atténuer l'impact de la formation antérieure et surtout de l'emploi durant la formation sur les compétences développées. Il s'agit par conséquent d'explorer la question de l'engagement dans la formation. Ce sont donc des pistes en matière d'ingénierie de formation et d'ingénierie pédagogique qui sont privilégiées à ce stade.

Altet, M. (2002). Entre réformes de l'école, rénovations des formations des enseignants et impact sur les pratiques enseignantes : décalages, passerelles et conditions. In M. Carbonneau & M. Tardif (Eds.), *Les réformes en éducation, leurs impacts sur l'école*. Sherbrooke : Editions du CRP, Université de Sherbrooke.

Astier, P. (2007). Alternance constituée, prescrite, vécue. *Education permanente*, 172, 61-72.

Gérard, F.-M. (2001). L'évaluation de la qualité des systèmes de formation. *Mesure et évaluation en éducation*, 24(2-3), 53-77.

Lambelet, D. (2009). Réinvestir le champ de la formation : éclipses et retour de la psychosociologie ? *Connexions*, 2009/2(92), 165-177.

Ourliac, G. (2014). Les résultats d'insertion sur le marché du travail peuvent-ils tenir lieu d'évaluation des politiques de formation ? In C. Beduwe, V. Bedin, & S. Croity-Belz (Eds.), *Evaluation formation emploi : un chantier pluridisciplinaire*. Paris: L'Harmattan.

Périsset, D., & Lehmann, L. (2010). Les recherches sur les formations à l'enseignement: perspectives et prospectives. *Revue suisse des sciences de l'éducation*, 32(1), 5-18.

Développement et implantation d'un programme par compétences en Kinésiologie : enjeux pédagogiques, évaluation et perspectives

Adrien Cantat, Francis Lalime, Philippe Corbeil, Geneviève Charland, Daniel Turpin

Le programme de Kinésiologie, rattaché à la faculté de médecine de l'Université Laval, a entamé son « Renouveau pédagogique » en 2013 à la suite à l'évaluation périodique du programme initié par le Vice-rectorat aux études et aux activités internationales. En réponse à cette évaluation, le programme a décidé de repenser complètement le design pédagogique du cursus en mettant l'accent sur l'approche programme et le développement de compétences. Ainsi, un programme articulé autour de six compétences à développer, que sont l'expertise clinique, la communication, la collaboration, la promotion de la santé, le professionnalisme et l'érudition, inspirées des rôles CanMEDS utilisés en médecine (Le Collège royal des médecins et chirurgiens du Canada, 2015), est en cours d'implantation depuis la session d'automne 2016. Un comité du « Renouveau pédagogique » a été créé à l'occasion pour orchestrer les travaux d'élaboration du nouveau programme.

Le développement d'un programme par compétences nécessite non seulement une refonte des méthodes d'enseignement et d'évaluation (Scallon, 2004), mais, également une réflexion profonde sur le profil souhaité de l'étudiant diplômé.

À défaut d'avoir un ordre professionnel auquel se référer pour avoir une ligne directrice en termes d'objectifs de formation tel qu'en médecine ou en physiothérapie, la première étape du « Renouveau

pédagogique » fut l'élaboration d'un référentiel de compétences. Le but de ce dernier est d'identifier les compétences nécessaires à la pratique de la profession de kinésologue (Paquay, 1994), puis de les détailler en composantes et sous-composantes. Dans le cadre du nouveau pédagogique, les étudiants devront ainsi faire état, tout au long du cursus, du niveau de développement des compétences ciblées et plus précisément des composantes définies.

Guidé par la réalisation d'une recension des écrits sur la question, le comité du Nouveau pédagogique a retenu après discussion plusieurs compétences qui ont été soumises pour validation aux enseignants ainsi qu'aux responsables et superviseurs de stages. Une fois les six compétences identifiées et définies, la même démarche de concertation et de validation a été réalisée à propos des composantes et des sous-composantes de chacune d'entre elles. Par la même occasion, cette étape a permis de dresser l'inventaire du contenu des cours du programme et d'établir concrètement leur lien avec le développement des compétences retenues tout en déterminant la gradation de la complexité des situations d'apprentissages. De fait, ce référentiel a initié l'instauration d'une approche programme permettant la consolidation des liens logiques entre les cours et favorisant la cohérence du programme par l'intégration progressive des apprentissages chez les étudiants (Sylvestre & Berthiaume, 2013).

Cette approche plus longitudinale implique par conséquent des modifications significatives dans l'architecture pédagogique du programme. Des changements ont été entrepris afin de privilégier le développement et l'évaluation des compétences des étudiants. De nouveaux cours longitudinaux intitulés « cours intégrateurs » (cinq cours) et « démarche clinique » (trois cours), répartis sur l'ensemble du cursus, seront insérés à chacune des trois années du programme. Leur rôle est primordial puisqu'ils visent le développement des compétences à travers la mobilisation, par les étudiants, de différentes ressources (Roegiers, 2000 ; Tardif, 2006) et la mise en application concrète des notions théoriques abordées dans les cours « traditionnels » davantage axés sur l'acquisition de connaissances. En outre, ces deux types de cours longitudinaux ont en commun l'exposition des étudiants à des situations d'apprentissage reflétant la réalité clinique de leur future profession.

Les « cours intégrateurs » ont pour objectif d'exposer les étudiants à une clientèle variée tout au long de leur cursus académique. Les interactions entre les étudiants et leurs clients se dérouleront sous la supervision directe du personnel enseignant et feront ainsi partie intégrante des activités d'apprentissages. Quant à eux, les cours « démarche clinique » sont avant tout des cours pratiques basés sur la réalité clinique des situations d'apprentissages. Ils visent essentiellement le développement des habiletés cliniques des étudiants à travers le raisonnement clinique, l'entrevue avec le client et l'interprétation des données découlant de la réalisation de l'examen physique (Ratté & Thériault, 2012). Pour cela, la démarche clinique se base sur une succession de six étapes méthodologiques distinctes composées au total de dix-neuf questions auxquels devront répondre les étudiants pour préciser, élaborer et justifier un plan d'intervention. Concrètement, celle-ci est opérationnalisée via la plateforme numérique DécliK visant à exposer les étudiants de manière fictive à une grande variété de clientèle.

Tout ceci a un impact certain sur les méthodes d'évaluation qui devront sortir de la dynamique traditionnelle « cours en classe et examen de connaissance ». En effet, des méthodes d'évaluation, comme des ECOS (Harden & Gleeson, 1979), des vignettes cliniques, des études de cas, permettant de porter un jugement sur les aspects techniques et cliniques ainsi que sur des compétences plus génériques (communication, professionnalisme), sont en cours d'élaboration.

Par ailleurs, celles-ci permettront également de documenter formellement le développement des compétences des étudiants en alimentant une plateforme numérique de suivi intitulée « Suivi du développement des compétences ». Tout au long des trois années d'études, chaque étudiant recevra régulièrement un tableau de suivi lui fournissant un portrait global du développement de ses compétences. En outre, dès le début du cursus, chaque étudiant sera associé à un tuteur avec qui il effectuera un suivi formel lors de rencontres planifiées. L'étudiant devra préalablement réaliser son autoévaluation avant d'en discuter avec son tuteur et de déterminer, sur la base du tableau de suivi, ou non des mesures d'amélioration (exercices à réaliser, exposition clinique à augmenter, etc.). Le programme de kinésiologie intègre ainsi dans son cursus un processus d'encadrement personnalisé des

étudiants.

Le passage à une approche par compétences implique immanquablement une refonte en profondeur du design pédagogique du programme que ce soit sur l'aspect pédagogique avec notamment la création de nouveaux cours, mais également en termes d'évaluation des apprentissages. La mise en place actuelle du « Renouveau pédagogique » dans le programme de kinésiologie a été un défi en termes de design et comporte encore des enjeux propres à la réinvention de certaines méthodes d'évaluation.

Bibliographie

Harden, R. M., & Gleeson, F. A. (1979). Assessment of clinical competence using an objective structured clinical examination (OSCE). *Medical Education*, 13, 39-54.

Le Collège royal des médecins et chirurgiens du Canada (2015). Référentiel de compétences CanMEDS 2015 pour les médecins. Extrait de <http://canmeds.royalcollege.ca/fr/referentiel>

Paquay, L. (1994). Vers un référentiel des compétences professionnelles de l'enseignant? *Recherche et Formation*, 16, 7-38.

Ratté, F., & Thériault, J. F. (2012). *Démarche clinique*. Québec, Canada : Presses de l'Université Laval.

Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Montréal, Canada : ERPI.

Sylvestre, E., & Berthiaume, D. (2013). Comment organiser un enseignement dans le cadre d'une approche-programme? Dans D. Berthiaume & N. Rege Colet (Eds.), *La pédagogie de l'enseignement supérieur: Tome 1: Enseigner au supérieur* (pp. 105-118). Berne, Suisse : Peter Lang.

Tardif, J. (2006). *L'évaluation des compétences. Documenter le parcours de développement*. Montréal, Canada : Chenelière Éducation.

Mots clés : évaluation, compétence, enseignement supérieur

Session 4 - Atelier 4.5

Jeudi 26 janvier 2017 – 10h30-12h00

Evaluation d'une plateforme internet d'apprentissage à l'écriture-lecture par des enseignants suivant une formation à son utilisation

Stéphanie Boéchat-Heer, Esther González-Martínez

Les technologies de l'information et de la communication (TIC) prennent de plus en plus de place à l'école suscitant nombre de questionnements en rapport aussi bien aux apprentissages des élèves qu'aux pratiques d'enseignement. En Suisse francophone, les médias et les nouvelles technologies font désormais partie intégrante des programmes d'enseignement et font l'objet de dispositifs de formation innovants. Toutefois, la recherche autant que l'expérience pratique montrent que l'intégration de ces dispositifs ne coule pas de source, les jeunes enseignants exprimant d'ailleurs autant de réticences que ceux plus expérimentés. L'intégration de ces dispositifs en classe dépend en fait de nombreux facteurs dont la qualité de la formation offerte aux enseignants futurs utilisateurs, les moyens technologiques et le soutien technique mis à leur disposition, le soutien de leurs collègues et supérieurs, et la culture des établissements scolaires en termes de technologie et d'innovation pédagogique (Boéchat-Heer, 2012; Fredricksson et Hoskins, 2007; Niemi et al. 2012; Valiente, 2010). Par ailleurs, nombreux travaux mettent en évidence l'existence d'un lien fort entre le sentiment d'auto-efficacité et l'acceptation de pratiques innovantes en matière d'intégration des TIC en éducation (Abbitt et Klett, 2007; Boéchat-Heer, 2011; Carugati et Tomasetto, 2002; Deaudelin, Dussault, Brodeur, 2002).

Le projet myMoment2.0 a été initié en 2005 par le Zentrum Lesen, Medien, Schrift de la Haute école pédagogique de Suisse occidentale (FHNW). Il a consisté à créer une plateforme internet sur laquelle des élèves peuvent écrire des histoires ainsi que lire, commenter et compléter celles de leurs camarades. Cette

plateforme vise à : a) développer la motivation à lire et à écrire, b) augmenter le niveau de lecture et d'écriture, c) développer des pratiques d'écriture coopérative, et d) apporter des compétences en matière de publication (dont des stratégies d'identification et attraction du public cible). Le centre interjurassien Médias, Images, Technologies de l'Information et de la Communication (MITIC) a adapté cette plateforme au français. En collaboration avec le centre MITIC, la Haute école pédagogique BEJUNE a mis en place une formation continue à l'attention des enseignants utilisant ou souhaitant utiliser myMoment2.0 dans des écoles du canton de Berne et Jura. Cette formation offre un ensemble d'enseignements et activités pédagogiques, mais aussi un accompagnement à l'utilisation de la plateforme et un espace de partage d'expériences entre experts et apprenants.

Cette communication présente une étude consacrée à l'évaluation que des enseignants se servant de cette plateforme font de celle-ci ainsi que de la formation continue les préparant à son utilisation. L'étude suit une démarche compréhensive de type qualitative (Kaufmann, 2011). Elle s'appuie sur la constitution et l'analyse d'un corpus d'entretiens semi-directifs auprès d'enseignants en français qui utilisent la plateforme avec des élèves âgés de 9 à 12 ans. Les entretiens ont fait l'objet d'une analyse thématique et catégorielle de contenu (Coffey et Atkinson, 1996) procédant en deux étapes : a) l'identification des thèmes centraux de chaque entretien et 2) l'identification des thèmes transversaux à l'ensemble du corpus (Miles et Huberman, 1994).

Les résultats de l'étude montrent l'importance du rapport entre la satisfaction des enseignants envers la plateforme et le déploiement de pratiques d'appropriation de celle-ci, par les enseignants et les élèves, adaptées aux circonstances pratiques et aux objectifs des utilisateurs. L'étude montre aussi que la satisfaction des enseignants avec la formation continue les préparant à l'utilisation de la plateforme dépend de la possibilité d'ajuster cette formation à leurs attentes, rythmes d'apprentissage, et niveaux de connaissances, autant que de la possibilité de s'en servir comme d'un espace de partage avec d'autres enseignants-utilisateurs. Notre communication présente ces résultats et, faisant écho aux thématiques du congrès, propose une réflexion sur le dispositif même d'évaluation et la manière de l'adapter pour qu'il informe au mieux le processus même de développement de la plateforme et de la formation continue qui lui est consacrée. Il sera notamment question des possibilités offertes par les TIC d'informer la conception et l'implémentation de dispositifs de formation au travers d'une analyse des pratiques d'utilisation alors que celles-ci sont en cours de déploiement.

Références

Abbitt, J. T. & Klett, M. D. (2007). Identifying influences on attitudes and self-efficacy beliefs towards technology integration among pre-service educators. *Electronic Journal for the Integration of Technology in Education*, 6. <http://ejite.isu.edu/Volume6/Abbitt.pdf> [On line]. Consulté le 19 octobre 2015.

Boéchat-Heer, S. (2011). Adaptation des enseignants aux usages des MITIC : sentiment d'auto-efficacité, formation et pratiques en classe. *Formation et pratiques d'enseignement en questions*, 12, 81-97.

Boéchat-Heer, S. (2012). Evaluation d'une formation aux TICE : développement de compétences et sentiment d'auto-efficacité. In S. Boéchat-Heer et B. Wentzel (Eds.), *Génération connectée: quels enjeux pour l'école?* (pp. 151-166). Bienne : Editions BEJUNE.

Carugati, F. & Tomasetto, C. (2002). Le corps enseignant face aux technologies de l'information et de la communication dans les pratiques d'enseignement. *Revue des sciences de l'éducation*, 28(2), 305-324.

Coffey, A., & Atkinson, P. (1996). *Making sense of qualitative data*. Thousand Oaks, CA : Sage.

Deudelin, C., Dussault, M. & Brodeur, M. (2002). Impact d'une stratégie d'intégration des TIC sur le sentiment d'auto-efficacité d'enseignants du primaire et leur processus d'adoption d'une innovation. *Revue des sciences de l'éducation*, 28(2), 391-410.

Fredriksson, U. and Hoskins, B. (2007). The development of learning to learn in a European context. *Curriculum Journal*, 18(2), 127-134.

Kaufmann, J.-C. (2011, 3ème Ed.). *L'entretien compréhensif*. Paris : Armand Colin.

Miles, M.B., & Huberman, A.M. (1994). *Analyse des données qualitatives*. Traduction de la 2e édition américaine (2005). Bruxelles : De Boeck.

Niemi, H., Kynäslahti, H., & Vahtivuori-Hänninen, S. (2012). Towards ICT in everyday life in Finnish schools: Seeking conditions for good practices. *Learning, Media and Technology*, 1, 1-15.

Valiente, O. (2010). 1-1 in Education. Current Practice, International Comparative Research Evidence and Policy Implications. OECD Education Working Papers, No. 44

Mots-clés : plateforme d'apprentissage ; lecture-écriture ; formation des enseignants

Recyclage de questions à choix multiples d'épreuves certificatives de physique en items de tests formatifs en ligne

Pierre-Xavier Marique, Jean-François Van De Poël, Maryse Hoebeke

Depuis septembre 2012, en Fédération Wallonie-Bruxelles de Belgique, les étudiants inscrits en première année de bachelier en médecine sont interrogés sous forme de questions à choix multiples lors des épreuves certificatives des cours scientifiques donnés au premier quadrimestre de leurs études. A l'Université de Liège, les examens du cours de physique sont organisés en deux parties composées chacune d'environ 25 questions. Les étudiants ont la possibilité de présenter cette épreuve à maximum 3 reprises. Chaque année académique, approximativement 150 questions inédites sont donc créées.

Ces questions sont répertoriées, et éventuellement améliorées, afin d'être réutilisées dans le cadre d'un outil d'entraînement en ligne, appelé « simulateur d'examens » (Marique, 2014, 2015). Cet outil est destiné à améliorer des connaissances directement en lien avec une stimulation constante de la mémoire des étudiants à l'aide de QCM (Roediger, 2006, 2011).

Même si lors de leur création, les 20 règles de conception développées par Leclercq (Leclercq, 1986) sont bien suivies, l'analyse a posteriori des questions d'examens permet leur amélioration en vue de leur introduction dans l'outil en ligne. A cet effet, chaque item est testé à l'aide de plusieurs indices.

- L'indice de difficulté P informe sur la proportion d'étudiants ayant correctement répondu à la question. Plus l'indice P est faible, plus la question sera jugée difficile. A l'inverse, plus l'indice P est élevé, plus la question est supposée facile. La valeur « corrigée » P', tenant compte d'un éventuel effet de hasard, est également calculée (Laveault & Grégoire, 1997).
- Le coefficient de discrimination D d'un item, servant à en déterminer la validité, est également mesuré. En effet, « un "bon" item est un item qui serait réussi par une plus grande proportion de sujets ayant obtenu un score élevé à l'examen que par des sujets ayant obtenu un score faible » (Laveault & Grégoire, 1997, p. 231). Autrement dit, ce coefficient devrait permettre de répondre à la double question suivante : les étudiants qui réussissent cet item sont également ceux qui ont obtenu un score élevé sur l'ensemble de l'épreuve ? (Kelley, 1939 ; Findley, 1956, Ebel ; 1965, Gilles, 2002) Les étudiants qui ont échoué à cet item font-ils également partie du groupe des scores les plus faibles sur l'ensemble de l'épreuve ? Pour ce faire, deux groupes sont isolés de la population totale : un premier groupe composé des 27 % des étudiants ayant obtenu les meilleures notes au test et un second reprenant les 27 % des étudiants caractérisés par les plus faibles notes. Le nombre d'étudiants ayant réussi l'item dans chaque groupe est alors mesuré. La différence divisée par le nombre d'individus de chaque groupe fournit la valeur du coefficient de discrimination.

En cas de valeurs non satisfaisantes pour ces deux indices, la question problématique est alors étudiée afin de déterminer ce qui a pu poser problème et éventuellement non retenue pour l'entraînement des étudiants. Après modification, elle est intégrée à l'outil en ligne. Les questions utilisées en ligne seront également ré-évaluées à intervalle régulier (1 fois par an).

L'ensemble des questions composant le simulateur d'examens est ensuite évalué par plusieurs membres de l'équipe pédagogique du Département de Physique de l'Université de Liège. Chacun attribue un niveau de difficulté à chaque question en fonction de 4 critères empiriques : abstraction, réflexion, mixité, mathématiques. Chaque évaluateur attribue une note de 1 à 3 à chacun des critères. La somme obtenue indiquera le niveau de difficulté total de la question selon la règle suivante : [4 ; 5] => 1 ; [6 ; 7] => 2 ; [8 ; 9]

=> 3 ; [10 ; 12] => 4.

Le critère d'abstraction est lié à la simplicité/complexité de l'énoncé, au fait que les données sont implicites/explicites, ... L'indice de réflexion est conditionné entre autres par le nombre d'étapes à effectuer pour parvenir à la solution, à la quantité de passage d'une formule à l'autre dans le raisonnement. La mixité sera simplement évaluée en fonction du nombre de matières, chapitres, compétences à mobiliser pour résoudre la question. Enfin, la valeur octroyée au critère mathématique dépendra de la quantité de formules et relations à utiliser, des compétences à mobiliser en algèbre et en analyse, ... L'attribution des valeurs pour chacun de ces critères est purement subjective. Cependant, afin d'assurer et de vérifier la concordance entre les évaluateurs, le kappa de Fleiss (Fleiss, 1981) a été mesuré pour chaque épreuve et nous indique, en général, une valeur approximative de 0,4, jugée satisfaisante.

Les questions sont alors classées par pool dépendant de leur niveau de difficulté et de la matière abordée. Elles permettront alors à l'étudiant un travail progressif sur le simulateur d'examens.

A l'avenir, dans une perspective d'amélioration du classement par niveau des questions disponibles dans le simulateur d'examens, la concordance entre l'indice P' d'un item, dépendant des résultats des étudiants à celui-ci, et le niveau de difficulté qui lui a été attribué par l'équipe pédagogique, sera étudiée.

L'impact de ce système sur l'apprentissage des étudiants fait actuellement l'objet d'un suivi particulier et précis. Toutes les tentatives des étudiants et leurs résultats sont enregistrés systématiquement en prévision d'une analyse des effets de ce système.

Références

BRENNAN, R.-L., *A generalized upper-lower item discrimination index*, Educational and psychological measurement, 32, 289-303, 1972

EBEL, R.L., *Confidence-Weighting and Test Reliability*, Journal of Educational Measurement, 2,49-57 B., 195

FINDLEY, W.G., *A rationale for evaluation of item discrimination statistics*, Educational and Psychological Measurement, 16, 175-180, 1956

FLEISS, J.-L., *Statistical methods for rates and proportions*, 2nd ed., John Wiley, New York, 1981

GILLES, J.-L., *Qualité spectrale des tests standardisés universitaires*, Thèse de doctorat, Université de Liège, 2002

KELLEY, T.L., *Selection of upper and lower groups for the validation of test items*, Journal of Educational Psychology, 30, 17-24, 1939

LAVEAULT, D., GREGOIRE, J. *Introduction aux théories des tests en sciences humaines*. Paris, Bruxelles : De Boeck Université & Larcier s.a., 1997

LECLERCQ, D. *La conception des questions à choix multiple*, Bruxelles, Labor, 1986

LECLERCQ, D. *Qualité des questions et signification des scores avec application aux QCM*, Bruxelles, Labor, 1987

MARIQUE, P.-X., HOEBEKE, M. *Plate-forme interactive au service des grandes populations d'étudiants suivant un cours de Physique*, Actes de la Conférence TICE 2014, Béziers, France, 2014

MARIQUE, P.-X., VAN DE POËL, J.-F., HOEBEKE, M. *Quel outil d'entraînement pour des étudiants en médecine évalués par QCM en physique ?* Actes du Colloque ADMEE 2015, Lisbonne, Portugal, 2015

ROEDIGER, H. L., BUTLER, A. C. *The critical role of retrieval practice in long-term retention*, Trends in Cognitive Sciences, volume 15, n°1, pp. 20–27, 2011. <http://dx.doi.org/10.1016/j.tics.2010.09.003>

ROEDIGER, H. L., KARPIPCKE, J. D., *The power of testing memory: Basic research and implications for educational practice*, Perspectives on Psychological Science, volume 1, n°3, pp. 181–210, 2006. <http://dx.doi.org/10.1111/j.1745-6916.2006.00012.x>

VIAU, R. *La motivation en contexte scolaire*, St-Laurent, Éditions du Renouveau pédagogique, 1994

https://www.uclouvain.be/cps/ucl/doc/edef/documents/EVA_QCM_version3.pdf, consulté le 24/09/2016

Exploration du développement de la compétence à évaluer des enseignants par la passation d'un test de jugement

Jennifer Smith, Eric Dionne

En 1996, le Rapport Delors intitulé « *L'éducation : un trésor caché dedans* » avait incontestablement eu un impact tangible sur les politiques de l'éducation à l'échelle mondiale en servant de cadre philosophique aux réformes de plusieurs pays, autant en Europe qu'en Amérique. Cette version renouvelée de l'apprentissage a mené de nombreux systèmes scolaires (Canada, Suisse, Australie, Grande-Bretagne, etc.) à apporter de profonds changements administratifs et pédagogiques afin de mieux préparer les jeunes aux défis qu'amène le 21^e siècle (Dionne, 2014). Plus spécifiquement, les approches contemporaines de l'éducation ont voulu faire face « à un monde en transformation rapide » et répondre à un nombre élevé de tensions suscitées par les changements technologiques, sociaux et économiques.

Ce changement paradigmatique issu du monde de l'éducation a nécessairement impliqué des ajustements majeurs, et ce, particulièrement sur le plan de l'évaluation des apprentissages. En effet, cette nouveauté a exigé du corps enseignant de revoir leurs pratiques évaluatives. Scallon (2004) explique que « la « méthodologie de l'évaluation a changé » puisqu'elle se caractérise maintenant par des tâches complexes, contextualisées et signifiantes permettant de mettre en œuvre plusieurs dimensions d'une compétence visée (De Ketele, 2000; De Ketele et Gérard, 2005; Figari, 2001; Paquay, 2002; Scallon, 2004; Roegiers, 2000; Tardif, 2006; Wiggins, 1998). Malgré les efforts du Ministère de l'époque à former son personnel enseignant en contexte changeant, on a dénoté les nombreuses « zones grises » liées au rôle d'évaluateur et le manque d'encadrement de la part des autorités quant à l'offre de formation continue en évaluation. Comme l'a mentionné Tardif (2006), l'évaluation des apprentissages est l'une des tâches les plus exigeantes du métier d'enseignant. Malgré tout, l'incertitude quant aux nouvelles pratiques évaluatives par compétence a représenté et représente toujours un grand défi (Perrenoud, 2001).

À la lumière de ces réalités, nous croyons qu'en s'inspirant de d'autres disciplines où le jugement professionnel est développé tout au long de la formation, les sciences de l'éducation pourraient ainsi faciliter le développement du jugement évaluatif par l'entremise d'outils formateurs mieux adaptés. Pour y arriver, il est devenu impératif de se pencher sur les dispositifs de formation au jugement afin d'en avoir une meilleure compréhension et ultimement d'arriver à les bonifier. De nombreux points de convergence ont été identifiés entre la profession médicale et celle du pédagogue, plus précisément entre le développement du jugement clinique et celui du jugement en évaluation. L'article de Jouquan, Parent et Audétat (2013) fait référence aux finalités centrées sur le développement de l'autonomie de la personne, la nature diagnostique de l'approche et de leurs objectifs ainsi qu'un parallélisme possible entre l'approche éducationnelle centrée sur l'apprenant et l'approche en santé centrée sur le patient. Inspirés par ces points de convergence, Dionne et Forte (2012) ont développé un instrument intitulé le test de jugement professionnel en évaluation des apprentissages (TJPEA). Composé de situations authentiques, le TJPEA vise à mesurer le niveau de jugement des étudiants à la formation des maîtres en contexte formatif. Afin d'y arriver, l'organisation de leurs connaissances est comparée à celle d'experts de la discipline. Les retombées pédagogiques positives soulignées par Dionne et Forte dans le cadre de leur expérimentation préliminaire du TJPEA de 2012 nous ont poussés à croire qu'il pourrait être pertinent de faire une expérimentation plus complète de l'instrument dans le contexte de notre étude.

L'objectif principal de notre collecte de données qualitative a été de mieux comprendre le processus de jugement en évaluation d'experts (n=4), de praticiens (n=4) et de novices (n=4) dans le cadre d'une passation orale d'un TJPEA. Le tout a été réalisé dans le cadre d'entrevues semi-dirigées d'une durée de 60 à 90 minutes enregistrées sur une bande audio. Les participants se sont vus présenter individuellement les quatre mêmes vignettes composées de 4 items. À chaque item, les participants ont partagé oralement leur opinion face aux contenus des situations présentées et ont justifié leurs actions. Le processus de collecte de données a permis l'analyse des différents discours. Une fois l'analyse verticale des contenus des entrevues de chacun des participants effectuée, nous avons effectué un croisement entre les discours de

celles-ci, et ce, au sein de chacun des groupes, puis entre les trois. Les similitudes et les divergences dénotées seront présentées plus explicitement dans le cadre de notre présentation.

En conclusion, depuis l'adoption de nouvelles pratiques en évaluation imposées par la réforme du système scolaire québécois des années quatre-vingt-dix, peu de recherches se sont intéressées à la formation nécessaire du corps enseignant à leur mise en œuvre. À ce jour, nous en savons très peu sur la façon dont les enseignants ont recours à divers processus mentaux lors de la prise de décisions liées au jugement en évaluation. Il devient alors fort difficile d'outiller les futurs enseignants qui seront appelés à évaluer quotidiennement dans le cadre de leurs fonctions. Même si nous sommes conscients que les résultats obtenus à cette présente étude ne sont pas généralisables en raison de son échantillon limité, il est tout de moins possible d'en faire découler de nouvelles avenues de recherche ainsi qu'une meilleure compréhension du processus de jugement évaluatif des enseignants. Cela nous permettra éventuellement de proposer des solutions à la relève enseignante et de favoriser le développement d'un jugement en évaluation juste et équitable pour les élèves.