



Association pour le développement des méthodologies d'évaluation en éducation @

Les visées plurielles de l'évaluation : Tensions et perspectives

34^e session d'études de l'ADMEE-Canada

15 et 16 novembre 2012

Hôtel Gouverneur Place Dupuis

1415, rue St-Hubert, Montréal (Québec) H2L 3Y9

☎ 1-888-910-1111 / Outre mer : 514-842-4881

Groupe **121115UMTL**

Mot de bienvenue des organisatrices,

C'est avec un immense plaisir que nous vous accueillons à la 34^e session annuelle de l'ADMEE-Canada, qui se tiendra les 15 et 16 novembre 2012 à l'Hôtel Gouverneur Place Dupuis à Montréal.

Cette année, nous avons choisi la grande métropole multiculturelle ouverte sur le monde qu'est Montréal pour nous réunir. Dans un élan de changement, nous avons également chamboulé les habitudes en vous proposant un programme comportant plus de présentations et enrichi d'une session d'affiches et d'une conférence de clôture. De plus, nous remettons deux prix d'excellence aux étudiants et nous ferons tirer des prix de présence. L'ajout d'une soirée festive sur les lieux du congrès, et à laquelle nous vous attendons très nombreux, clôturera notre première journée.

Les sessions des années passées se sont intéressées aux aspects politiques, méthodologiques, éthiques et professionnels de l'évaluation. Pour 2012, nous avons retenu un thème englobant, « Les visées plurielles de l'évaluation, tensions et perspectives », puisqu'il fait référence à la fois aux diverses utilisations de l'évaluation dans des visées diagnostique, formative et certificative et aux multiples objets propres à être évalués en éducation que sont les politiques et les systèmes éducatifs, ainsi que leurs composantes.

Cette session d'études offre une excellente occasion d'apporter des éléments de réponse à de multiples questions reliées aux changements qui ont pris place dans le monde de l'éducation depuis une dizaine d'années. Ainsi, les congressistes, qu'ils soient chercheurs, praticiens, ou professionnels et intervenants des différents milieux d'application de l'évaluation, pourront échanger et débattre sur des sujets aussi variés que l'évaluation de programmes, l'analyse des pratiques des formateurs ou des résultats des élèves que ce soit au Canada, en Europe, en Afrique, en Amérique latine, mais aussi sur les défis méthodologiques sous-jacents à leur étude.

Nous avons choisi un lieu de rencontre au cœur de la ville de Montréal qui nous accueillera autant pour les moments scientifiques que festifs tout au long de ces deux journées. L'Hôtel Gouverneur Place Dupuis, situé à proximité du métro Berri-UQAM, nous offre un tarif préférentiel et un stationnement intégré pour permettre à chacun et chacune de loger sur place s'il le souhaite et ainsi privilégier les échanges pendant toute la durée de cette session d'études.

Il nous reste à vous rappeler que vous êtes les artisans du succès de ce congrès qui est un évènement majeur du monde de l'évaluation et à vous remercier d'avoir été si nombreux à répondre à notre invitation.

Au nom de l'ADMEE-Canada nous vous souhaitons à tous la bienvenue et un congrès fructueux.

Les organisatrices,
Nathalie Loye
Micheline-Joanne Durand

Les visées plurielles de l'évaluation : Tensions et perspectives

Horaire de la 34^e session d'études

Jeudi 15 novembre 2012	09 h 00 – 10 h 00	Accueil et inscription
	10 h 15 – 11 h 45	Conférence d'ouverture
	11 h 45 – 13 h 15	Dîner libre
	13 h 15 – 13 h 45	Communications Bloc A
	14 h 00 – 14 h 30	Communications Bloc B
	14 h 45 – 15 h 15	Communications Bloc C
	15 h 30 – 16 h 00	Communications Bloc D
	16 h 15 – 16 h 45	Communications Bloc E
	17 h 00 – 19 h 00	Session d'affiches
	19 h 00 – 20 h 30	Souper (personnes déjà inscrites)
	20 h 30 – 22 h 00	Soirée festive (pour tous / Bar payant)

Vendredi 16 novembre 2012	08 h 30 – 09 h 00	Inscription
	09 h 00 – 10 h 30	Ateliers 1 et 2
	10 h 30 – 10 h 45	Pause
	10 h 45 – 11 h 15	Communications Bloc F
	11 h 30 – 12 h 00	Communications Bloc G
	12 h 00 – 13 h 45	Réunion et dîner de l'ADMEE
	13 h 45 – 14 h 15	Communications Bloc H
	14 h 30 – 15 h 00	Communications Bloc I
	15 h 15 – 15 h 45	Communications Bloc J
	16 h 00 – 16 h 30	Conférence de clôture
	16 h 30 – 17 h 00	Remise des prix

Jeudi 15 novembre 2012**Conférence d'ouverture (de 10 h 15 à 11 h 45)****Réflexions sur le déploiement de la quantification et de la mesure en éducation**Jean-Guy **Blais** (Université de Montréal)**BLOC A (Jeudi de 13 h 15 à 13 h 45)**

Jeudi 13 h 15 ⊗	A1	Étude des annotations d'un enseignant à la suite de l'enseignement explicite des stratégies de résolution de problèmes mathématiques Joanne Lapointe et Micheline-Joanne Durand (Université de Montréal)
Jeudi 13 h 15 ⊗	A2	Estimation des paramètres du modèle de Rasch à l'aide des logiciels R et BILOG : une étude comparative Sébastien Béland et Gilles Raïche (UQAM)
Jeudi 13 h 15 ⊗	A3	Développement et validation de grilles descriptives holistiques pour l'évaluation de la résolution de problèmes en mécanique pour ingénieurs Maxim Morin (Université de Montréal), Lina Forest , Thomas Gervais et David Ménard (École Polytechnique de Montréal)
Jeudi 13 h 15	A4	Sélection, prédiction et compétences: réflexion sur la validité des mesures et critères Gilles Leclerc (Université de Montréal)

BLOC B (Jeudi de 14 h à 14 h 30)

Jeudi 14 h	B1	Stratégies d'évaluation formative non instrumentées selon différents contextes d'enseignement: les classes régulières, de maternelle et d'éducation physique Joëlle Morrisette et Grégoire Compaoré (Université de Montréal)
Jeudi 14 h	B2	Validation d'un questionnaire diagnostique du français à l'aide de la théorie de la réponse aux items Francis Djibo (Cégep de Rivière-du-Loup)
Jeudi 14 h	B3	Évaluation automatisée de compétences technologiques et informationnelles des étudiants accédant aux universités québécoises Michel Pronovost , Jean Allard , Robert Gérin-Lajoie (Université de Montréal), Annie Mahais Kowaliszyn , Yves Munn et Richard Guay
Jeudi 14 h	B4	Évaluation de l'enseignement par les étudiantes et étudiants : un modèle hiérarchique et une comparaison des modalités de passation en ligne vs papier Léon Harvey et Marie-Hélène Hébert (UQAR)

BLOC C (Jeudi de 14 h 45 à 15 h 15)

Jeudi 14 h 45	C1	Analyse de la complexité de situations évaluatives de compétences Chantal Roussel (UQAR)
Jeudi 14 h 45	C2	Étude de la modélisation des liens entre certaines variables associées à l'environnement scolaire et la performance à l'écrit des élèves Patrick Charles et Jean-Guy Blais (Université de Montréal)

Jeudi 14 h 45	C3	CONFÉRENCE ANNULÉE
Jeudi 14 h 45	C4	Nouvelle méthodologie d'évaluation des programmes à l'UCL : de la résistance à l'adhésion, du singulier au collectif, d'une vue administrative à une démarche qualité Jean-Marc Braibant et Christine Jacqmot (Université catholique de Louvain, Belgique)

BLOC D (Jeudi de 15 h 30 à 16 h)

Jeudi 15 h 30 ⊗	D1	SimulaCAT : un progiciel pédagogique facilitant la compréhension du testing adaptatif Christian Bourassa et Gilles Raïche (UQAM)
Jeudi 15 h 30	D2	Validation pronostique de l'évaluation d'une compétence en recourant à la théorie des ensembles flous Paul Martin
Jeudi 15 h 30 ⊗	D3	Croyances et pratiques liées à l'évaluation de l'oral réflexif : étude de cas de deux enseignants de français de 3^e secondaire Myriam Vega (Université de Montréal)
Jeudi 15 h 30 ⊗	D4	L'évaluation diagnostique de la compréhension en lecture : besoins et perspectives Catherine Croisetière et Gilles Raïche (UQAM)

BLOC E (Jeudi de 16 h 15 à 16 h 45)

Jeudi 16 h 15	E1	Gestion des visées, tensions et perspectives de l'évaluation : point de vue du superviseur clinique lors de l'évaluation de la performance Christina St-Onge, Martine Chamberland et Annie Lévesque (Université de Sherbrooke)
Jeudi 16 h 15	E2	L'assurance qualité des apprentissages dans les facultés et les écoles d'administration des affaires, le rôle de la validité Sawsen Lakhali, Éric Frenette et Serge Sévigny (Université Laval)
Jeudi 16 h 15	E3	Construction d'un outil diagnostique en orthographe française destiné aux futurs maîtres Pierre Paradis, Lorraine Pepin, Marie-Hélène Hébert et André Jacques (UQAR)
Jeudi 16 h 15 ⊗	E4	Les théories du domaine de l'évaluation de programme : une cartographie Sylvain Dubé et Nathalie Loye (Université de Montréal)

PRÉSENTATIONS PAR AFFICHE (Jeudi de 17 h – 19 h 00)

Affiche 1 ⊗	Les principales difficultés rencontrées par les enseignants de français au secondaire lors de l'évaluation des lectures interprétatives Alexie Miquelon, Sylvie Marcotte et Sébastien Béland (UQAM)
Affiche 2 ⊗	L'évaluation certificative en fin du primaire au Sénégal : de l'indépendance à nos jours Ibrahima Sakho et Micheline-Joanne Durand (Université de Montréal)
Affiche 3 ⊗	Mesurer la maîtrise de l'information à l'université : un projet pilote québécois Catherine Séguin et Nathalie Loye (Université de Montréal)
Affiche 4 ⊗	Lien entre l'évaluation des apprentissages et la motivation des étudiants : un tour d'horizon à l'aide de la métaanalyse

-
- Nadine **Talbot** et Gilles **Raïche** (UQAM)
- Affiche 5** **Les effets des grandes politiques d'évaluation sur le curriculum et l'enseignement au Brésil**
⊗ Andressa **Aita Ivo** (Université Laval)
- Affiche 6** **Évaluation des établissements et réussite éducative : et si l'enjeu réel était l'évaluation des enseignants?**
⊗ Nora **Larfi** (Université de Montréal)
- Affiche 7** **ANNULÉE**
- Affiche 8** **L'évaluation diagnostique cognitive des compétences en langues – Exemples de quelques modèles de mesures utilisés dans le contexte des tests à grande échelle**
⊗ Dan **Thanh Duong Thi** et Nathalie **Loye** (Université de Montréal)
- Affiche 9** **Pertinence des épreuves du MELS et rendement des élèves**
⊗ France **Paquette** et Micheline-Joanne **Durand** (Université de Montréal)
- Affiche 10** **Documenter les pratiques évaluatives des enseignants de la 6^e année du primaire**
⊗ Mireille **Mouffe**, Caroline **Leblanc** et Micheline-Joanne **Durand** (Université de Montréal)
- Affiche 11** **Évaluer le raisonnement clinique avec un test de concordance de script : un exemple en sciences infirmières**
⊗ Marie-Ève **Latreille** et Éric **Dionne** (Université d'Ottawa)
- Affiche 12** **Analyse de la qualité des épreuves ministérielles de 6^e année de français et de mathématiques de 2010, 2011 et 2012**
⊗ Marie-Hélène **Asselin** et Micheline-Joanne **Durand** (Université de Montréal)
- Affiche 13** **Indicateurs de qualité de questions écrites : élaboration et premières étapes de validation d'une grille de vérification**
⊗ Jacinthe **Cloutier**, Yann **Tissier**, Sébastien **Joncas**, André-Sébastien **Aubin** et Christina **St-Onge** (Université de Sherbrooke)

19 h 00 **Souper** Personnes inscrites d'avance (65\$) sur présentation d'un coupon

20 h 30 – 22 h 00 **Soirée Festive** Pour tous (Service de bar payant / Musique)

Vendredi 16 novembre 2012

ATELIERS 1 et 2 (Vendredi de 9 h à 10 h 30)

Vendredi 9 h	Atelier 1	Démystifier les tests ontariens à grande échelle de l'Office de la qualité et de la responsabilité en éducation Micheline Wilkie (OQRE/Office de la qualité et de la responsabilité en éducation)
Vendredi 9 h	Atelier 2	Table ronde sur l'évaluation des compétences professionnelles dans une approche programme à la formation des maîtres au Québec Pascale Lefrançois (Université de Montréal), Alain Rousseau (UQTR) et Sylvie Viola (UQAM)

BLOC F (Vendredi de 10 h 45 à 11 h 15)

Vendredi 10 h 45	F1	Diagnostiquer les besoins des élèves et choisir le bon outil, un gage de réussite Julie Provencher (MELS)
Vendredi 10 h 45 ⊗	F2	Analyse comparative de la dimensionnalité d'un test de placement en anglais langue étrangère selon la TRI multidimensionnelle et la méthode bifactorielle Christophe Chénier (UQAM)
Vendredi 10 h 45	F3	Oser certifier l'acquisition de compétences par des examens à choix multiple : l'approche de la Commission de la construction du Québec Jonathan Bélair , Anne-Andrée Bourdeau , Laila Valin et Francis Kelly (CCQ)
Vendredi 10 h 45 ⊗	F4	La contribution de l'évaluateur aux processus de production et de crédibilisation du jugement Marie-Pier Marchand (UQAC) et Marthe Hurteau (UQAM)

BLOC G (Vendredi de 11 h 30 à 12 h)

Vendredi 11 h 30	G1	Évaluer la langue dans différents pays : une étude comparative d'épreuves ministérielles Pascale Lefrançois (Université de Montréal) et Catherine Brissaud (Université de Grenoble)
Vendredi 11 h 30	G2	CONFÉRENCE ANNULÉE
Vendredi 11 h 30	G3	Une culture scolaire d'évaluation qui comprend les tests provinciaux en Ontario Hélène Boudreau (OQRE/EQAO, Ontario)
Vendredi 11 h 30	G4	Validité de l'évaluation en apprentissage par cas : comparaison entre l'évaluation automatisée basée sur l'argumentation et l'évaluation manuelle Geneviève Gauthier (Université de l'Alberta)

BLOC H (Vendredi de 13 h 45 à 14 h 15)

Vendredi 13 h 45	H1	La méthodologie : le parent pauvre du rapport d'évaluation de programme? Éric Dionne (Université d'Ottawa)
Vendredi 13h45	H2	Analyse d'un dispositif de formation à l'évaluation des apprentissages à la lumière du cadre de la validité globale

		Réjean Auger (OgDDASt inc.) et Nicole Landry (UQTR)
Vendredi 13 h 45 ⊗	H3	Analyse de la valeur pédagogique et recherche évaluative : évaluation d'outils pédagogiques Judith Beaulieu et Jacques Langevin (Université de Montréal)
Vendredi 13 h 45 ⊗	H4	L'analyse de la satisfaction professionnelle chez les nouveaux enseignants et les chevronnés en utilisant la modélisation de Rasch Carla Barroso da Costa et Nathalie Loye (Université de Montréal)

BLOC I (Vendredi de 14 h 30 à 15 h)

Vendredi 14 h 30	I1	L'examen de l'impact de la position d'un item dans un plan matriciel multiple sur sa difficulté estimée : cas du Programme international pour le suivi des acquis des élèves (PISA) Koffi-Philippe Houme (Conseil des ministres de l'éducation du Canada) et Yunmei Xu (OQRE/EQAO, Ontario)
Vendredi 14 h 30 ⊗	I2	Mesurer et évaluer les compétences et la capacité à innover Pierre Gignac et Nathalie Loye (Université de Montréal)
Vendredi 14 h 30 ⊗	I3	La formation initiale des maitres et la didactique du français oral : regard sur les grilles d'observation de l'oral Annabelle Caron , Isabelle Carignan et Vincent Grenon (Université de Sherbrooke)
Vendredi 14 h 30 ⊗	I4	Le portfolio professionnel : un outil d'évaluation dans la formation des maitres Hélène Meunier , Gilles Raïche et Carole Raby (UQAM)

BLOC J (Vendredi de 15 h 15 à 15 h 45)

Vendredi 15 h 15	J1	Élaborer des grilles d'évaluation à échelles descriptives France Côté (Cégep Marie-Victorin) et Hélène Meunier (UQAM)
Vendredi 15 h 15	J2	Étude comparative de la préparation scolaire des enfants à la maternelle en milieux défavorisés montréalais Pierre Lapointe (Université de Montréal)
Vendredi 15 h 15	J3	CONFÉRENCE ANNULÉE
Vendredi 15 h 15	J4	La conception de tests d'évaluation en lecture, construction de questions et de corrigés Éric Falardeau , Pierre Valois et Frédéric Guay (Université Laval)

Conférence de clôture (Vendredi de 16 h à 16 h 30)

Tendances et visée plurielles en évaluation

Michel **Laurier** (Université d'Ottawa)

Remise des prix :

- Prix d'excellence meilleure présentation par affiche étudiante ⊗
- Prix d'excellence meilleure présentation orale étudiante ⊗
- Prix de présence

Jeudi 15 novembre 2012

Conférence d'ouverture (Jeudi de 10 h 15 à 11 h 45)

Réflexions sur le déploiement de la quantification et de la mesure en éducation

Jean-Guy Blais (Université de Montréal)

En ce début de 21^e siècle, la pratique de la mesure dans les sciences sociales est toujours largement tributaire d'habitudes, d'idées, de concepts, de modes, de préjugés qui ont jalonné le développement de toutes les disciplines scientifiques depuis le siècle des lumières. Mais, alors que la mesure et la quantification l'ont nettement emporté dans les sciences de la nature, elles ont rencontré des terrains d'application dans les sciences sociales qui engendrent des difficultés mettant en relief une réussite toute relative. La nature même des objets d'étude respectifs illustre bien les difficultés éprouvées. Par exemple, à quand remonte la dernière fois où un électron a refusé de participer à une expérience, a tenté de dissimuler la vérité, n'a pas fait preuve de motivation ou est né dans un milieu défavorisé?

Il n'y a rien de surprenant à ce qu'on veuille donner la forme de la quantité ou de la mesure aux données récoltées en éducation car, de tout temps, la mesure a favorisé la convergence des regards et de l'interprétation. Ainsi, de nos jours tout ce qui est présenté sous la forme de nombres se voit rapidement attribuer les propriétés de la mesure car cette transposition ouvre la porte à un schème d'interprétation commun et les interlocuteurs ont alors l'impression de parler la même langue, de se comprendre. Alors qu'à d'autres époques une citation latine ou une éloquence soignée rassuraient sur le sérieux d'un individu et la confiance qu'on pouvait lui témoigner, de nos jours c'est le nombre (les mesures et les statistiques) et la précision qui lui est automatiquement attribuée (le plus souvent sans preuve) qui marquent la rhétorique du discours de celui qui veut en imposer à ses concitoyens. Beaucoup de gens semblent en effet faire plus confiance à la personne qui cite des chiffres à l'appui de ce qu'elle avance qu'à celle qui ne le fait pas.

Mais ce n'est pas parce qu'on utilise des outils mathématiques ou statistiques sophistiqués que l'objet d'application de ces outils se révèle posséder *de facto* les propriétés de ce qui fait une quantité. La quantité n'est pas une propriété innée que toutes les caractéristiques de l'humain ou de la société possèdent par définition. Ainsi, si nous réclamons le statut de quantités et de mesures pour les données que nous recueillons, il faut aussi réfléchir aux avenues qui s'offrent à nous pour montrer que ces données possèdent les attributs de la quantité. En bref, il faut s'intéresser à la mesure et à son système de production.

Jean-Guy Blais est professeur depuis 1990 au département d'administration et fondements de l'éducation à l'Université de Montréal. Il est titulaire d'un baccalauréat en mathématiques (1978), d'une maîtrise en statistique (1980) et d'une doctorat en mesure et évaluation (1987). Il a également effectué un stage postdoctoral d'une année (1987-1988) au *Educational Testing Service* à Princeton aux États-Unis. Il dirige actuellement une équipe de recherche interuniversitaire (GRIÉMÉtic; <http://www.griemetic.ca/accueil.aspx>) subventionnée par le Fonds québécois pour la recherche sur la société et la culture (FQRSC) qui étudie les systèmes, technologies et modèles de mesure pour l'évaluation en éducation/formation. Ses intérêts de recherche sont l'histoire et la philosophie de la mesure dans les sciences sociales, les critères d'application des modèles de mesure avec les tests et les questionnaires, la validité des données pour la recherche et l'évaluation en éducation, l'accord inter-juges, l'évaluation des apprentissages avec les technologies de l'information et de la communication.

BLOC A (Jeudi de 13 h 15 à 13 h 45)

Jeudi 13 h 15

A1 Étude des annotations d'un enseignant à la suite de l'enseignement explicite des stratégies de résolution de problèmes mathématiques

Joanne Lapointe et Micheline-Joanne Durand (Université de Montréal)

Le Programme de formation de l'école québécoise situe l'élève au cœur de ses apprentissages. L'enseignant peut faciliter le développement des compétences en offrant une rétroaction permettant à l'élève de progresser dans ses apprentissages. Il est difficile pour les enseignants de faire des annotations pertinentes et efficaces en

mathématique, car l'accent est mis sur le concept travaillé et non sur la démarche mathématique. C'est pourquoi, nous avons porté notre regard sur l'incidence que peuvent avoir l'enseignement explicite des stratégies ainsi que les annotations faites par l'enseignant sur les copies des élèves en ce qui a trait au développement de leurs compétences à résoudre des problèmes complexes en mathématique.

Nous avons opté pour une recherche qualitative et collaborative pour vivre un échange avec l'enseignant et une interinfluence entre le praticien et le chercheur. La technique d'échantillonnage retenue pour le choix de l'enseignant a été celle de cas exemplaires, tandis que celle que nous avons choisie pour les élèves était l'échantillonnage intentionnel critérié. La recherche a duré du mois de novembre au mois de mai de l'année scolaire 2008-2009. Comme instruments de cueillette de données, nous avons opté pour des entrevues avec l'enseignant et des mini-entrevues avec les élèves à deux moments de la recherche. Nous avons consulté les travaux corrigés des élèves dans leur portfolio.

Notre étude fait ressortir l'apport de l'enseignement stratégique de la démarche mathématique. Les résultats précisent que les annotations de type méthodologique ont été celles qui ont été les plus utilisées et ont permis une meilleure compréhension chez l'élève. De plus, elles favorisent le transfert d'une situation à l'autre et permettent à l'élève d'obtenir de meilleurs résultats.

Jeudi 13 h 15

A2 Estimation des paramètres du modèle de Rasch à l'aide des logiciels R et BILOG : une étude comparative

Sébastien **Béland** et Gilles **Raïche** (UQAM)

La théorie de la réponse aux items (TRI) est un modèle de mesure fort utilisé en éducation. À ce jour, de nombreux logiciels, tel que BILOG, sont disponibles afin de procéder à l'estimation des paramètres d'items et de calculer les scores des sujets à l'aide de cette modélisation. Parmi ces logiciels, R n'est pas à négliger car, il présente de nombreux avantages : il est gratuit, est fort documenté et présente de nombreuses librairies permettant de faire des analyses très variées.

Cette communication a comme objectif de comparer l'estimation des paramètres de sujets et d'items selon une des modélisations issues de la TRI, le modèle de Rasch, à l'aide de BILOG et de plusieurs librairies disponibles dans le logiciel R. À cette fin, nous allons utiliser les données d'un test de classement en anglais, langue seconde. Les résultats seront comparés sur la base de l'erreur type et des paramètres d'items et de sujets estimés par BILOG, comme base de comparaison, et différentes librairies de R.

Jeudi 13 h 15

A3 Développement et validation de grilles descriptives holistiques pour l'évaluation de la résolution de problèmes en mécanique pour ingénieurs

Maxim **Morin** (Université de Montréal), Lina **Forest**, Thomas **Gervais** et David **Ménard** (École Polytechnique de Montréal)

Au Canada, avant 2015, les facultés et écoles de génie auront la responsabilité de conformer leurs programmes de premier cycle à une nouvelle norme établie par le Bureau canadien d'agrément des programmes de génie. Elles devront notamment attester que les diplômés possèdent douze qualités fondamentales à l'exercice de la profession d'ingénieur.

Le cours de Mécanique pour ingénieurs, offert à l'École Polytechnique de Montréal à 550 étudiants par année, est un exemple typique de cours qui devra intégrer la maîtrise des douze qualités dans son cadre d'évaluation, en mettant l'accent sur deux d'entre elles, soit les connaissances en sciences et génie, et l'analyse de problèmes. L'évaluation des apprentissages dans ce cours porte alors sur l'analyse et la résolution de problèmes avec applications numériques, ce qui est concrètement réalisé au moyen d'examens composés de questions à réponses construites. La notation des solutions se

fait traditionnellement en décernant des points partiels aux éléments de réponses absents ou présents, mais il s'avère difficile dans ces conditions d'interpréter de façon juste et équitable la diversité des démarches et des réponses produites par les étudiants (Henderson, Yerushalmi, Kuo, Heller et Heller, 2004).

L'objet de cette communication consiste à exposer les résultats de la première phase d'une étude consacrée à l'évaluation de résolution de problèmes en mécanique. Il sera d'abord question des résultats de l'analyse des besoins et d'une recension des écrits qui ont conduit à une classification des problèmes d'examen ainsi qu'à l'élaboration de grilles descriptives holistiques. Dans un second temps, les résultats de la validation de ces grilles auprès d'experts de contenu et de correcteurs expérimentés seront discutés. Finalement, la présentation fera état des défis et des retombées de cette démarche pour la planification de l'apprentissage et de l'évaluation dans ce cours.

Jeudi 13 h 15

A4 Sélection, prédiction et compétences : réflexion sur la validité des mesures et critères
Gilles **Leclerc** (Université de Montréal)

Traditionnellement, pour une institution universitaire ou un organisme professionnel, les démarches de sélection des candidats reposent principalement sur un classement centré sur les cotes de rendement. Il est cependant, légitime de se questionner sur la validité de ces données pour la sélection des futurs professionnels. Ces cotes de rendement reflètent-elles la complexité et la diversité des compétences professionnelles requises à l'exercice de la profession? Il est reconnu que le virage du paradigme des connaissances vers celui des compétences demande des transformations significatives, tant dans la culture professionnelle que dans les dynamiques pédagogiques et les procédures d'évaluation. Dans un tel paradigme, la conception de l'apprentissage et la portée de l'enseignement se transforment. L'évaluation devient un processus intégré à même l'apprentissage qui valorise l'authenticité des situations et des contextes d'apprentissage. Un tel virage soulève la critique quant aux fonctions de l'évaluation à privilégier et suscite la réflexion quant aux objets ou construits à évaluer (manifestations observables). De ce fait, ce paradigme remet en question les indicateurs de l'apprentissage qui se doivent de faire dorénavant état du niveau de développement différencié des compétences. En regard de cette évolution qu'adoptent progressivement les programmes professionnels universitaires, les processus de sélection doivent-ils s'ouvrir au paradigme des compétences? À cet effet, l'importance des enjeux sociaux et des décisions académiques demande que tant les données utilisées pour la sélection des étudiants que les critères de performance définis pour classer et décrire le développement professionnel s'avèrent valides et cohérents avec la démarche professionnelle. La présentation propose, alors, une réflexion, à la fois théorique et pragmatique, sur la validité des démarches de sélection des étudiants et de prédiction des compétences en contexte universitaire.

BLOC B (Jeudi de 14 h à 14 h 30)

Jeudi 14 h

B1 Stratégies d'évaluation formative non instrumentées selon différents contextes d'enseignement : les classes régulières, de maternelle et d'éducation physique

Joëlle **Morrisette** et Grégoire **Compaoré** (Université de Montréal)

Avec l'avènement de la réforme en cours au primaire depuis 2001, les enseignantes et les enseignants sont invités à modifier leurs pratiques évaluatives dans le sens d'une utilisation quotidienne de la fonction formative, nécessaire au développement des compétences (Legendre, 2004). L'évaluation formative renvoie au cumul et à l'interprétation d'information sur les apprentissages des élèves, recueillie au cours

même des activités d'enseignement/apprentissage, visant à identifier leurs forces et faiblesses, à les aider à réfléchir à leurs processus d'apprentissage et à guider leurs démarches pour soutenir leur progression et, enfin, à orienter la planification de l'enseignement (Andrade et Cizek, 2009). Cette fonction de l'évaluation, qui peut revêtir différentes formes, implique par conséquent deux processus : la construction d'un jugement sur les apprentissages réalisés par chacun, de même qu'une intervention visant à soutenir la progression de ces apprentissages (Morrissette, 2010a).

Jusqu'à présent, plusieurs ont documenté les stratégies instrumentées d'évaluation formative, telles les pratiques de testing et de contrôle (Campos et O'Hern, 2007), les portfolios (Mottier Lopez et Allal, 2004), etc., mais peu ont investigué celles mises en œuvre de façon plus informelles, qui se construisent dans l'interaction en classe, dans des processus de coconstruction de sens (par exemple, Bell et Cowie, 2001; Leug et Mohan, 2004; Morrissette, Mottier Lopez et Tessaro, à paraître), et donc se rapportant à des observations et des discussions en classe.

Dans le cadre de cette contribution, nous jetterons une lumière sur les stratégies d'évaluation formative non instrumentées déployées par des enseignantes qui sont en appropriation face à la réforme (FQRSC, 2010-2013). Nous chercherons à mettre en exergue pourquoi certaines stratégies sont privilégiées selon les différents contextes d'enseignement, en distinguant les manières de faire utilisées en classe régulière, de maternelle.

Jeudi 14 h

B2 Validation d'un questionnaire diagnostique du français à l'aide de la théorie de la réponse aux items

Francis **Djibo** (Cégep de Rivière-du-Loup)

Plusieurs tests diagnostiques du français langue première existent dans le réseau collégial sans qu'il soit possible de confirmer leurs qualités métrologiques. Cette recherche est une contribution visant à relever ce défi. Le cégep de Rivière-du-Loup développe un test diagnostique dont le but est d'offrir de la rétroaction à des étudiants faibles et au personnel du cégep sur leur progression dans la maîtrise du français.

Une calibration du test sous forme de questions à choix multiples se fait à l'aide d'un modèle à deux paramètres affilié à la théorie de la réponse aux items (TRI). Les données statistiques sont produites à l'aide du logiciel R dont certaines fonctions donnant plus de contrôle pour le calcul des coefficients d'ajustement sont présentées.

Les résultats préliminaires sur les coefficients de fidélité, sur le processus de sélection des items selon les paramètres du modèle et sur les cas de fonctionnement différentiel sont présentés. Une ouverture est faite sur le traitement des cas de réponses inappropriées.

Jeudi 14 h

B3 Évaluation automatisée de compétences technologiques et informationnelles des étudiants accédant aux universités québécoises

Michel **Pronovost**, Jean **Allard**, Robert **Gérin-Lajoie** (Université de Montréal), Annie **Mahais Kowalyszyn**, Yves **Munn** et Richard **Guay**

Les pratiques actuelles d'évaluation des compétences technologiques et informationnelles nécessaires à la poursuite d'études universitaires nécessitent des ressources importantes pour la correction manuelle des évaluations. Les tests automatisés qui sont présentement disponibles au Québec ne mesurent pas la connaissance ou la compétence, mais la perception qu'on a de notre connaissance ou comportent des questions de connaissances, qui ne mesurent pas la compétence ou la capacité à mettre en pratique et à résoudre des problèmes.

Le projet Certitude, projet inter-ordre réalisé par l'Université de Montréal, le Collège Jean-de-Brébeuf et CÉGEP à distance, propose de mesurer les habiletés pour ces deux

domaines et vise à mettre en place un environnement automatisé d'évaluation des compétences.

Le projet Certitude a réalisé une revue des référentiels de compétence informationnelle et technologique, colligé et analysé quelques instruments d'évaluation existants, réalisé un survol des types de technologies utilisées dans ces instruments. Il a recensé trois classes de méthodologie et technologie d'évaluation automatisées : les évaluations basées sur des questions à choix multiples (QCM), les évaluations basées sur des QCM « augmentées » par du multimédia et de la simulation générique, les évaluations basées sur des plugiciels de simulation insérés dans les outils technologiques. Nous avons développé deux nouveaux types d'évaluation *in situ* : l'une avec le tableur qui génère des jetons et l'autre qui analyse les balises XML de fichiers de type traitement de texte.

Le projet a retenu une approche combinant ces différents types de questions (questions à choix multiple augmentées, simulation, réalisation « *in situ* » et « chasse au trésor ») pour les différentes habiletés visées : traitement de texte, tableur, recherche d'information, traitement de l'image et environnement de travail.

Jeudi 14 h

B4 Évaluation de l'enseignement par les étudiantes et étudiants : un modèle hiérarchique et une comparaison des modalités de passation en ligne vs papier

Léon Harvey et Marie-Hélène Hébert (UQAR)

L'évaluation de l'enseignement par les étudiants (EEE) désigne les procédures par lesquelles les étudiants prennent part au jugement porté sur des cours ou des enseignements spécifiques, attribués à des professeures et professeurs au sein d'un programme ou d'une formation. Les enjeux sont importants. Les résultats sont utilisés à des fins académiques et formatives, mais sont également évoqués d'une manière sommative lors de décisions liées à la carrière professorale. Dans le cadre de la présente recherche, deux objectifs sont poursuivis. Premièrement, la documentation est divisée sur la nature du construit qui doit être mesuré. S'il existe une multitude d'instruments valides et fiables, ceux-ci mesurent des construits multidimensionnels différents, ce qui rend variable le processus de mesure d'une institution à l'autre. Ces instruments sont constitués d'une part d'items spécifiques qui mesurent des dimensions particulières souvent très corrélées et d'items généraux qui sont des indicateurs d'un jugement global sur l'enseignement. Notre premier objectif de recherche est donc de déterminer si un modèle hiérarchique ou encore de second ordre (analyse factorielle secondaire) permettent de mieux rendre compte de la dualité des items qu'un modèle factoriel de premier ordre. Deuxièmement, avec l'avènement des technologies de l'information et de la communication, les conditions dans lesquelles se déroulent ces évaluations ont été modifiées. Notre second objectif est de vérifier les impacts de ces changements sur la validité et la fiabilité des évaluations. Le corpus de données est constitué des évaluations réalisées durant les années 2007-2008 (20 245 évaluations) et 2010-2011 (16 432 évaluations) à l'Université du Québec à Rimouski. Les résultats attestent qu'un modèle hiérarchique décrit bien la structuration des évaluations, qu'un instrument valide et fiable est utilisé et que le taux de réponses est affecté par la modalité en ligne. Les implications sont discutées.

BLOC C (Jeudi de 14 h 45 à 15 h 15)

Jeudi 14 h 45

C1 Analyse de la complexité de situations évaluatives de compétences

Chantal Roussel (UQAR)

Depuis quelques décennies, le concept de compétence se décline en de multiples

définitions. Même si l'on observe une volonté de consensus entourant ce concept, il subsiste plusieurs zones floues. Suivant chacune des définitions observées, la compréhension et l'interprétation de ce que représente une compétence orientent le choix, la conception et la mise en œuvre de situations permettant le développement et l'évaluation des compétences dans les programmes d'études. Les travaux de recherche que nous poursuivons établissent d'abord la relation entre la définition de compétence privilégiée et ses caractéristiques, en nous inspirant des travaux de Tardif (2006). C'est cette compréhension du concept de compétence que nous transposons ensuite dans une architecture particulière. Nous pouvons alors, découlant de cette structure, établir ce qui caractérise les situations évaluatives en termes de ressources, d'exigence intégrative et de contextualisation. En lien étroit avec les éléments sélectionnés, les écrits de Roegiers (2007) concernant la notion de complexité que l'on peut associer aux situations évaluatives employées pour évaluer des compétences, cette notion se voit également utilisée, puis modulée en prenant en compte la définition et les caractéristiques retenues pour la compétence. À partir de cette conceptualisation, nous analysons à l'aide de différents continuums la complexité présente dans les situations évaluatives employées en formation professionnelle au secondaire. Nous examinons d'abord les ressources (nombre et variété) présentes, l'articulation privilégiée entre ces ressources (exigence de la tâche), de même que le type de contexte dans lequel se déroule la situation évaluative. L'analyse de ces données nous invite à questionner la compréhension actuelle des situations évaluatives utilisées dans une perspective d'évaluation des compétences, puisqu'il semble exister différents écarts entre les propositions théoriques et les pratiques ayant cours actuellement.

Jeudi 14 h 45

C2 Étude de la modélisation des liens entre certaines variables associées à l'environnement scolaire et la performance à l'écrit des élèvesPatrick **Charles** et Jean-Guy **Blais** (Université de Montréal)

Une étude du MELS rapporte que les élèves de 5^e secondaire éprouvent de la difficulté à obtenir de bons résultats à deux critères de correction (syntaxe et ponctuation, orthographe), parmi les six de l'épreuve unique de français écrit. Ce fait nous amène à nous intéresser à l'étude de la modélisation des liens entre des facteurs associés à l'environnement scolaire et les résultats des élèves de la région de Montréal, en ce qui a trait à ces deux critères spécifiques pour les cohortes des années 2006, 2007 et 2008.

Nous procédons d'abord à des analyses descriptives des variables pour chacune des trois populations. Nous poursuivons l'analyse en effectuant plusieurs modélisations multiniveaux multivariées des deux critères en fonction de variables indépendantes caractéristiques de l'élève et de l'école.

Les résultats indiquent une relative stabilité dans la performance des élèves pour les trois années, avec une légère amélioration de la performance pour la cohorte de 2007, et qui se maintient pour la cohorte de 2008. Les élèves du secteur privé obtiennent de meilleurs résultats que ceux du public. Le résultat des filles est supérieur à celui des garçons et les élèves de langue maternelle française obtiennent de meilleurs résultats que ceux de langues maternelles différentes du français. Il importe cependant d'apporter quelques nuances dans l'interprétation de ces résultats.

La part de variance dans la performance des élèves attribuable à l'élève est de l'ordre de 75 % et de l'ordre de 25 % à l'école. On note un effet différentiel des facteurs élève et école selon le critère considéré. Les variables caractéristiques de l'élève comptent pour 13,9 % de la variance totale de la performance des élèves en syntaxe et ponctuation et pour 9,8 % en orthographe. Les variables caractéristiques de l'école comptent pour 3,7 % de la variance totale de la performance des élèves en syntaxe et ponctuation et pour 6,4 % en orthographe.

Jeudi 14 h 45

C4 Nouvelle méthodologie d'évaluation des programmes à l'UCL : de la résistance à l'adhésion, du singulier au collectif, d'une vue administrative à une démarche qualitéJean-Marc **Braibant** et Christine **Jacqmot** (Université catholique de Louvain, Belgique)

Trop souvent, l'évaluation se limite à un recueil de données, alors que cette phase devrait constituer la première étape d'un processus qui en comporte au moins trois (Dory, 2012) :

- le recueil,
- la comparaison des données à une référence afin d'émettre un jugement (conforme ou non aux attentes),
- la prise de décision.

Notre présentation se focalisera sur une approche mise en œuvre depuis peu à l'Université catholique de Louvain (UCL) afin de développer la seconde étape et de soutenir la troisième.

Depuis quelques années, l'UCL évalue ses programmes au moyen d'un questionnaire centré sur la formation des étudiants et sur les effets des formations dispensées au sein des programmes de BAC et de Master. L'outil s'inspire des travaux de P. Ramsden ainsi que des items dérivés de cette recherche par l'ITL de l'Université de Sydney (Jacqmot et coll., 2009).

Cet outil a été adapté pour inclure une méthodologie permettant la lecture commentée et contrastée des résultats à partir d'un diagnostic des points forts et des points faibles de chacun des programmes évalués. Très vite, le besoin de se donner des points de référence afin de situer le programme est apparu. Nous avons donc développé une approche statistiquement validée permettant une mise en regard du programme cible par rapport aux autres programmes de l'UCL. Cela a littéralement « boosté » la discussion collégiale en termes d'analyse, de regards croisés et de prises de décisions. Cette constatation ainsi que le nombre croissant de demandes qui nous parviennent laissent à penser que notre approche soutient le développement d'un véritable pilotage collégial des programmes, dans une approche plus formative qu'administrative, en exploitant un cadre de référence comme un espace de liberté et non pas comme une contrainte. Cette approche est aisément transférable en enseignement supérieur.

BLOC D (Jeudi de 15 h 30 à 16 h)

Jeudi 15 h 30

D1 SimulaCAT : un progiciel pédagogique facilitant la compréhension du *testing* adaptatifChristian **Bourassa** et Gilles **Raïche** (UQAM)

Bien que largement utilisées depuis des décennies, la théorie classique des tests et l'administration de tests identiques quelles que soient les personnes ne sont pas toujours les meilleures stratégies pour estimer le niveau d'habileté d'une personne. Les modélisations issues de la théorie de la réponse à l'item offrent des avantages notables par rapport à la théorie classique des tests. Notamment, celles-ci facilitent l'analyse du fonctionnement différentiel d'items, la détection de patrons de réponses inappropriés ou encore, ce qui nous intéressera ici, la mise en application de tests adaptatifs.

Le *testing* adaptatif est l'une des applications dont l'apparition a été facilitée par le développement des modélisations issues de la théorie de la réponse à l'item. Il permet l'administration d'un test sur mesure pour un individu en ciblant les items les plus informatifs sur son niveau d'habileté. Ainsi, après chaque item répondu, un niveau d'habileté provisoire est estimé pour l'individu et le procédé se répète jusqu'à ce qu'un

certain nombre d'items ait été administré ou qu'une certaine précision de mesure soit atteinte.

Le progiciel SimulaCAT a été développé dans une perspective pédagogique. Plus spécifiquement, il a été conçu afin de faciliter la compréhension générale du *testing* adaptatif et de ses spécificités à l'aide de simulations de passation. Notamment, il produit des rapports de passation et plusieurs graphiques pour comprendre son fonctionnement sur le plan technique : courbes d'information, courbes de vraisemblance, courbes caractéristiques d'item, etc. SimulaCAT a été développé dans un environnement Microsoft Access et fait appel à la librairie catR de R.

Cette communication vise, d'une part, à familiariser les participants au *testing* adaptatif à l'aide de SimulaCAT et, d'autre part, à expliquer son fonctionnement.

Jeudi 15 h 30

D2 Validation pronostique de l'évaluation d'une compétence en recourant à la théorie des ensembles flous

Paul Martin

La question de la validité des épreuves pour l'évaluation des compétences demeure très pertinente. Que l'on pense à la compétence « savoir écrire des textes variés » que l'on évalue à la fin du secondaire et au collégial. Comment se fait-il que l'on accuse encore les étudiants diplômés de ces niveaux de ne pas savoir écrire? Pire, des facultés universitaires et des employeurs font passer des épreuves en français tant ils ne font pas confiance aux épreuves qui évaluent cette compétence. Donc, on se doit de prouver la validité pronostique d'un instrument qui évalue une compétence pour rétablir ce lien de confiance.

La méthode classique de prouver la validité pronostique d'une épreuve consiste à calculer la corrélation entre deux ensembles de scores. Or, ramener à un simple score l'évaluation d'une compétence signifie simplifier ce qui est complexe; on perd alors l'authenticité de l'évaluation. Par contre, la validité pronostique d'une épreuve peut être abordée sous un autre angle. En effet, une démonstration avec la théorie des ensembles flous a déjà été effectuée avec une épreuve de mathématique (l'auteur, dans Raïche et al., 2011). La théorie des ensembles flous a été développée en 1965 par Lofti Zadeh et a subi par la suite des raffinements pour trouver de nombreuses applications dans des domaines variés dont celui de l'éducation.

Dans le but d'appliquer cette nouvelle méthode de validité pronostique, on aura recours à une simulation avec les résultats à l'épreuve de fin de secondaire pour la compétence « savoir écrire des textes variés » et la même épreuve au collégial. L'hypothèse est que les élèves qui savaient écrire à la fin du secondaire, savent aussi écrire lors de l'épreuve au collégial.

Raïche, G., Paquette-Côté, K. et Magis, D. (2011). Des mécanismes pour assurer la validité de l'interprétation de la mesure en éducation : Volume 1 - La mesure. Éd. PUQ, p. 121-134.

Jeudi 15 h 30

D3 Croyances et pratiques liées à l'évaluation de l'oral réflexif : étude de cas de deux enseignants de français de 3^e secondaire

Myriam Vega (Université de Montréal)

Des études récentes signalent que l'oral est très peu enseigné dans les classes de français langue d'enseignement et que son évaluation est une source d'inconfort pour les enseignants (Lafontaine et Messier, 2009; Sénéchal, 2012). Bien que recommandées par le programme de formation actuel, les activités d'oral réflexif en interaction, comme les cercles littéraires, sont également peu pratiquées en classe et les enseignants ignorent comment évaluer la progression des élèves dans ce contexte (Hébert, 2010).

Dans le cadre d'une recherche exploratoire portant sur la lecture littéraire et l'oral

réflexif (Hébert et Lafontaine, FQRSC-MELS, 2009-2013), nous avons suivi deux enseignants de français de troisième secondaire afin de documenter leurs croyances et pratiques déclarées ayant trait à l'évaluation de l'oral et de l'oral en interaction, ainsi que leurs pratiques observées d'évaluation de l'oral aux étapes de la prise d'information, de l'interprétation et du jugement (Laurier, Tousignant et Morissette, 2005).

L'activité évaluative serait le résultat d'une combinaison de dimensions (culture de l'évaluation, postures de l'évaluateur, compétence évaluative, geste évaluatif) qui, de par leur complexité et leur conjugaison, seraient susceptibles de créer des tensions chez les enseignants (Jorro, 2009). Ainsi, sachant qu'un changement de paradigme en évaluation des apprentissages a été adopté avec l'arrivée des programmes axés sur le développement de compétences en éducation (Tardif, 2006), nous avons cherché à savoir de quelle façon ces différentes dimensions se sont manifestées au cours du projet et quelles formes de contradictions, de ruptures ou d'incompatibilités ont émergé dans ce contexte d'évaluation de la compétence orale.

La présente communication fera état de nos résultats de recherche, en ce qui a trait aux pratiques d'évaluation de l'oral réflexif en contexte de cercles littéraires entre pairs, telle que vécue par les enseignants participants.

Jeudi 15 h 30

D4

L'évaluation diagnostique de la compréhension en lecture : besoins et perspectives

Catherine Croisetière et Gilles Raïche (UQAM)

À l'aube des années 2000, l'utilisation d'instruments d'évaluation fiables et valides en classe est impératif. Or, alors qu'une majorité d'enseignants est d'avis que le recours à des évaluations formatives est le meilleur moyen de collecter une information diagnostique à propos des élèves (Huff et Goodman, 2007), les études démontrent que celle-ci n'est pas toujours bien construite. Ainsi, cette situation peut poser problème si le résultat obtenu n'est pas suffisamment précis, si le jugement porté n'est pas valide et si elle engendre des interventions inadéquates.

Cette communication, issue de la réalisation d'une thèse doctorale, vise trois principaux objectifs. D'abord, elle cherche à démontrer la nécessité, tant pour les chercheurs que pour les praticiens, de développer les instruments d'évaluation diagnostiques, qui peuvent faciliter grandement les apprentissages (De la Torre, 2009). Ce développement est d'ailleurs particulièrement souhaitable en langue, la maîtrise de cette dernière étant à la base du succès dans toutes les autres disciplines (Taboada, Tonks, Wigfield et Guthrie, 2008). Dans un deuxième temps, cette communication fait état des forces et lacunes des quelques instruments d'évaluation diagnostique actuellement disponibles en compréhension en lecture en français pour les élèves du primaire (Cormier, Desrochers et Sénéchal, 2006; Jang, 2009), dont *l'Évaluation diagnostique de la lecture* (Armand et Sabourin, 1995), les *Indicateurs dynamiques des savoirs essentiels en lecture* (St-Laurent et Giasson, 2009) et *Le vol du P.C.* (Boutard, Claire et Gretchanovsky, 2010). Enfin, la présentation se termine en indiquant les principales caractéristiques devant guider la conception des prochains instruments d'évaluation diagnostique de la compréhension en lecture. Plus précisément, les étapes à suivre, la forme (type d'items, tâche, type de texte, etc.) et le contenu (sujet, vocabulaire, processus et connaissances, etc.) des instruments sont discutés.

BLOC E (Jeudi de 16 h 15 à 16 h 45)

- Jeudi 16 h 15 E1 **Gestion des visées, tensions et perspectives de l'évaluation : point de vue du superviseur clinique lors de l'évaluation de la performance**
Christina **St-Onge**, Martine **Chamberland** et Annie **Lévesque** (Université de Sherbrooke)
- Introduction : Les écrits scientifiques portant sur l'évaluation de la performance indiquent que les évaluateurs représentent une source d'erreur de mesure non négligeable. Néanmoins, ils ont un rôle important dans le processus d'évaluation, devant observer, interpréter et juger la performance de leurs étudiants et résidents. La participation des évaluateurs demeure cependant incontournable dans le processus d'évaluation de la performance clinique. Par conséquent, une meilleure compréhension de leurs processus d'évaluation pourrait permettre une modélisation des meilleures pratiques et ainsi augmenter la qualité de l'évaluation de la performance clinique.
- Objectif : L'objectif de cette étude était de documenter les visées, tensions et perspectives à propos de l'évaluation que possèdent des superviseurs cliniques lors de l'évaluation de la performance clinique d'une résidente.
- Méthodologie : Des entrevues semi-structurées ont été réalisées avec 11 évaluateurs (nommés pour leur excellence) d'un département de médecine afin de faire émerger les visées, tensions et perspectives pendant l'évaluation d'une performance clinique via vidéo. Une analyse basée sur les principes de la théorisation ancrée a été effectuée.
- Résultats : Trois thèmes ont émergé de l'analyse des données, 1- les sources extrinsèques d'information, 2- les attentes intrinsèques des évaluateurs, et 3- les processus cognitifs sollicités pour rallier ces sources d'information.
- Discussion : L'interaction des attentes externes vs les attentes de l'évaluateur a mis en évidence la gestion des visées et tensions sous-jacente à cette tâche. Lorsque la performance observée différait de la performance attendue, les évaluateurs prenaient un recul afin de réexaminer leur interprétation et leur jugement de la performance. Cette stratégie, un processus décisionnel double (*dual-processing*), devrait faire l'objet d'études futures.
- Jeudi 16 h 15 E2 **L'assurance qualité des apprentissages dans les facultés et les écoles d'administration des affaires, le rôle de la validité**
Sawsen **Lakhal**, Éric **Frenette** et Serge **Sévigny** (Université Laval)
- The *Association to Advance Collegiate Schools of Business*, organisme américain d'agrément des facultés d'administration des affaires, a mis en place des standards d'assurance qualité des apprentissages, nommés *Assurance of Learning Standards*. Ces standards incitent à la mise en place d'une approche programme en utilisant une démarche en cinq étapes itératives. La première étape consiste à établir un profil de sortie par programme menant à un grade universitaire. Ce profil de sortie doit regrouper un certain nombre d'énoncés décrivant les résultats attendus des apprentissages des étudiants, communément appelés *Learning Outcomes*. Ces résultats des apprentissages, énoncés sous forme de buts, doivent refléter le degré d'atteinte des compétences acquises dans les programmes et attendues sur le marché du travail. La deuxième étape revient à aligner les curriculums avec les buts des programmes. Elle consiste à vérifier jusqu'à quel point les programmes déjà existants couvrent convenablement les buts tels que définis. En ce qui concerne la troisième étape, les professeurs doivent mesurer jusqu'à quel point les étudiants ont réellement acquis ces buts de programmes en utilisant des méthodes d'évaluation diverses et appropriées. Pour ce faire, ils doivent mettre en place des outils d'évaluation ainsi que

des seuils de performance. La quatrième étape consiste à analyser les données. Enfin, en ce qui concerne la cinquième étape, les données de l'évaluation doivent être utilisées pour l'amélioration continue des programmes. Lors de cette présentation, il sera question de l'influence de cette démarche itérative en cinq étapes sur la validité des résultats obtenus quant aux apprentissages des étudiants.

Jeudi 16 h 15

E3 Construction d'un outil diagnostique en orthographe française destiné aux futurs maîtres

Pierre **Paradis**, Lorraine **Pepin**, Marie-Hélène **Hébert** et André **Jacques** (UQAR)

Depuis 2003, les fautes de français écrit d'étudiants en sciences de l'éducation de l'UQAR sont colligées de façon systématique dans leurs travaux personnels dans le cadre de cours du programme de formation des maîtres. Plus de 5 000 fautes ont, à ce jour, été analysées par les membres de notre équipe de recherche. Les fautes ont été classées dans une grille comprenant les grandes catégories (p. ex. accord du participe passé avec « avoir ») et sous-catégories (p. ex. complément direct inexistant, situé après le verbe, situé devant le verbe) en orthographe grammaticale et d'usage.

Partant de ce corpus, nous avons construit un outil diagnostique informatisé dont l'originalité tient au fait qu'on retrouve dans celui-ci les mêmes proportions des différents types de fautes que celles observées dans l'établissement du portrait de l'ensemble des étudiants. Parce que ces fautes représentent l'état de la situation de nos étudiants après leur passage dans une formation secondaire puis collégiale, nous émettons l'hypothèse que le Test diagnostique des compétences en orthographe française (Paradis, Pepin, Hébert et Jacques, 2012) évalue les difficultés réelles des étudiants dans leurs écrits quotidiens. La tâche des usagers consiste à détecter et à corriger des fautes d'orthographe grammaticale et d'usage. Les résultats donnés à l'utilisateur au sortir du test sont à la fois quantitatifs (on lui donne une note, sur 100) et qualitatifs (on lui fait connaître les catégories et sous-catégories qui ont été plus ou moins bien réussies).

L'objet de la communication est de mettre au jour le test diagnostique évoqué plus haut et d'exposer le processus qui a conduit à sa construction, ainsi que sa contribution souhaitée au développement de la compétence des futurs maîtres à communiquer à l'écrit dans la langue d'enseignement.

Jeudi 16 h 15

E4 Les théories du domaine de l'évaluation de programme : une cartographie

Sylvain **Dubé** et Nathalie **Loye** (Université de Montréal)

L'histoire de l'évaluation de programme montre que sa nature est changeante et contingente. Chacune des périodes et générations de ce domaine diffère par la fonction attribuée à l'évaluation, le rôle joué par les évaluateurs, le degré d'intérêt manifesté par les acteurs sociaux, notamment. Les contextes scientifique, politique et économique ont en permanence joué un rôle critique et déterminé l'éventail des moyens disponibles pour l'évaluation. Plusieurs chercheurs éminents du domaine de l'évaluation de programme vont même jusqu'à dire que les impératifs politiques et sociaux ont constitué des moteurs de changements plus importants que les développements des théories en évaluation eux-mêmes (Stufflebeam, 2007; King 2003). Or, caractériser les théories du domaine de l'évaluation de programme n'a jamais été une chose aisée, ce qui semble constituer une difficulté inhérente à l'évaluation tant elle est complexe à la fois dans son fonctionnement et dans ses applications. Nous référer à une « théorie de l'évaluation », en admettant qu'une telle théorie est fondée et justifiée, s'avère inapproprié tant la notion même de théorie de l'évaluation pose problème. Dans le cadre de cette communication, nous proposons une cartographie des diverses théories de l'évaluation de programme. Cette approche s'est avérée nécessaire étant donné l'hétérogénéité des éléments que nous avons

retenus pour cette communication : théorie prescriptives et descriptives, métathéories, approches et modèles d'évaluation de programmes.

Présentations par affiche (Jeudi de 17 h à 19 h)

Affiche 1 **Les principales difficultés rencontrées par les enseignants de français au secondaire lors de l'évaluation des lectures interprétatives**

Alexie **Miquelon**, Sylvie **Marcotte** et Sébastien **Béland** (UQAM)

« Comprendre et interpréter un texte » est une composante de la compétence à atteindre en lecture à l'école secondaire québécoise (MELS, 2007). En classe, cependant, l'interprétation est beaucoup moins travaillée que la compréhension (Beaudry, 2009; Tauveron, 2001), notamment parce qu'elle est mal définie et mal comprise de la part des enseignants (Falardeau, 2003) et que la lecture interprétative pose des problèmes dans le cadre de son évaluation. En effet, contrairement à sa compréhension, l'interprétation d'un texte est difficilement évaluable, en raison de « l'activité subjective du lecteur » (Dufays, Gemenne et Ledur, 2005, p. 322).

Cette communication par affiche présentera les résultats de deux études menées à partir de cette problématique. D'abord, les résultats d'une recension des écrits sur les principaux problèmes en évaluation des lectures interprétatives, tels que répertoriés par des chercheurs en didactique de la lecture, seront présentés sous forme d'une taxonomie que nous avons élaborée. Ensuite, seront présentés les résultats d'un questionnaire passé à une vingtaine d'enseignants du secondaire et permettant de mettre en exergue les difficultés qu'ils rencontrent lors de l'évaluation de la composante « interprétation » de la compétence en lecture. Cette seconde étude aura pour but de confirmer la validité de la taxonomie développée à partir de la recherche théorique menée préalablement.

Affiche 2 **L'évaluation certificative en fin du primaire au Sénégal : de l'indépendance à nos jours**

Ibrahima **Sakho** et Micheline-Joanne **Durand** (Université de Montréal)

Après son accession à la souveraineté internationale en 1960, le Sénégal amorce une tentative de nationalisation de son système éducatif. Depuis, il n'a cessé d'évoluer et la dernière réforme jusqu'ici s'inscrit dans la mouvance mondiale d'un nouveau curriculaire qui a touché plusieurs pays occidentaux : l'approche par compétences. Dans la foulée de ces réformes, l'évaluation certificative en fin du cycle primaire, instituée bien avant l'indépendance, connaîtra une évolution compte tenu d'une dynamique complexe qui s'opère au sein d'un imbroglio sociopolitique qui mettait en présence plusieurs acteurs. Autrement dit, la nouvelle vision politique devrait tenir compte de l'environnement socioculturel, ce qui n'était pas évident à l'époque. En effet, les premières années d'indépendance sont marquées par la volonté du nouvel État indépendant de rompre avec le système à tendance assimilationniste pour adopter un système à tendance émancipatrice. L'option fut de mettre l'accent sur les disciplines dites fondamentales comme le français et les mathématiques, ce qui semblait être contradictoire sachant que le principal véhicule de la culture est la langue (le français, langue seconde étant la langue d'enseignement dans le pays). Plus tard, cette évaluation certificative subira quelques modifications par suite des conclusions de différentes structures consultatives.

Aujourd'hui, l'évaluation certificative continue sa transformation. La communication proposée analyse l'évolution du système d'évaluation certificative en fin du primaire au Sénégal sous l'angle de la justice, de l'égalité et de l'équité. Pour ce faire, nous passerons en revue les grands moments de l'évolution de l'évaluation certificative au primaire au Sénégal à partir de 1960 et soulèverons les principaux enjeux qui ont accompagné chaque type d'évaluation. Également, nous examinerons les perspectives envisageables pour la nouvelle orientation relative à l'évaluation des compétences.

Affiche 3 **Mesurer la maîtrise de l'information à l'université : un projet pilote québécois**

Catherine **Séguin** et Nathalie **Loye** (Université de Montréal)

La maîtrise de l'information est un levier et un enjeu dans la société. Par ailleurs, l'université

québécoise tarde à se l'approprier (Séguin, 2012). Certains auteurs avancent l'hypothèse que le manque de données empiriques pour démontrer la contribution de la maîtrise de l'information à divers enjeux sociétaux ne permet pas de convaincre les acteurs du domaine (Hinchliff, 2005; Katz, 2007).

Cette communication vise, dans un premier temps, à faire un tour d'horizon des instruments et procédures utilisés dans diverses universités anglo-saxonnes pour mesurer le niveau de maîtrise de l'information de leurs étudiants, suivi par une réflexion sur leur applicabilité en contexte québécois.

Dans un deuxième temps, nous présenterons un instrument sous forme de questionnaire actuellement en développement en contexte francophone québécois à l'UQO. Nous décrirons le cadre théorique, le questionnaire qui en découle, ainsi que le mode de passation mis en place en septembre 2012. Nous présenterons les premiers résultats issus des analyses descriptives des réponses obtenues, ainsi que ceux issus de l'application d'un modèle de Rasch afin d'étudier les propriétés métriques de notre instrument.

Affiche 4 Lien entre l'évaluation des apprentissages et la motivation des étudiants : un tour d'horizon à l'aide de la métaanalyse

Nadine **Talbot** et Gilles **Raïche** (UQAM)

À la lecture d'études menées sur le lien entre l'évaluation des apprentissages en salle de classe et la motivation des étudiants, force est de constater que les instruments de mesure utilisés ainsi que les méthodes d'analyses de données diffèrent. Cette différence ne permet pas nécessairement de préciser si le lien est fort ou faible. À ce moment, le lecteur éprouve de la difficulté à comparer les résultats de ces études et ainsi ne peut se faire une idée globale des articles dont il a pris connaissance. Afin de rendre comparable ces résultats, il est possible de transformer les résultats statistiques en une mesure statistique uniforme, appelée taille de l'effet, pour toutes les études. Notamment, il peut s'agir du coefficient de corrélation, du *d* de Cohen ou du *g* de Hedges. Cette façon de faire fait partie de la démarche d'une métaanalyse.

Au cours de notre exposé, les résultats d'études sur le lien entre l'évaluation des apprentissages en salle de classe et la motivation des étudiants seront présentés, notamment en utilisant le coefficient de corrélation qui s'avère une solution de rechange appropriée.

Affiche 5 Les effets des grandes politiques d'évaluation sur le curriculum et l'enseignement au Brésil

Andressa **Aita Ivo** (Université Laval)

Les réformes éducatives qui ont été initiées à partir des années 1990 au Brésil ont entraîné de profonds changements notamment sur les plans de la gestion, du travail enseignant et du curriculum. Cet article a pour but d'analyser plus particulièrement les effets des politiques d'évaluation sur le curriculum, compte tenu qu'elles mettent l'accent sur la centralisation et la régulation dans la perspective du nouveau management basé sur le rendement (Hypolito, 2010; Garcia et Anada, 2009; Oliviera, 2004). Cette recherche fait partie d'un projet en cours dans le réseau scolaire de la municipalité de Santa Maria, au Brésil. L'étude a été réalisée, selon une approche qualitative, dans deux écoles du réseau municipal d'enseignement et auprès du Secrétariat municipal de l'éducation (SMED). Les données ont été recueillies à partir de deux sources. La première consiste à faire l'analyse documentaire des lois éducatives municipales, des règlements qui en découlent concernant le curriculum et des curricula en vigueur dans les écoles. La seconde source de données provient des entrevues semi-structurées (N = 13) qui furent réalisées auprès des membres de l'équipe administrative et pédagogique (3) du SMED, de professeurs (3) et de membres (2) de l'équipe de gestion de chacune des écoles. Les résultats des analyses permettent de faire ressortir les effets importants des politiques d'évaluation sur les pratiques pédagogiques et les curricula des écoles. De telles évaluations proposent une vision de l'apprentissage déterminée par des examens standardisés qui ne prennent pas suffisamment en compte la spécificité du contexte scolaire, les particularités des élèves et de leur évolution dans le processus d'enseignement/apprentissage. De plus, cette politique cherche à rendre les enseignants responsables des apprentissages des élèves mais en même temps exempte les administrateurs de leurs propres responsabilités dans ce processus.

Affiche 6 Évaluation des établissements et réussite éducative : et si l'enjeu réel était l'évaluation des enseignants?**Nora Larfi (Université de Montréal)**

La question de l'évaluation des établissements s'est imposée dans la foulée de la réforme des programmes de l'école québécoise et de l'administration publique en général. La gestion axée sur les résultats a fait son entrée dans le monde de l'éducation en introduisant des changements dans la Loi sur l'instruction publique (LIP). Pour accompagner ces changements, des outils de gestion se sont imposés à tous les échelons : un plan stratégique et une Convention de partenariat, qui relie chacune des commissions scolaires au Ministère de l'Éducation, du Loisir et du Sport (MELS); une Convention de gestion qui engage chaque commission avec chacun des établissements relevant de son territoire; et, à l'échelle de chaque établissement, un plan de réussite et un projet éducatif qui sont implantés.

Ces outils sont conçus dans le but de répercuter la prise en charge des 5 buts du MELS : (1) Augmentation du taux de diplomation et de qualification avant l'âge de 20 ans (2) Amélioration de la maîtrise de la langue française; (3) Amélioration de la persévérance scolaire et de la réussite de certains groupes cibles, comme EHDA; (4) Amélioration de l'environnement sain et sécuritaire dans les établissements; (5) Augmentation du nombre d'élèves de moins de 20 ans en formation professionnelle.

Cibler des objectifs chiffrés et mesurables est au cœur de ce nouveau mode de gestion qui exige la mobilisation de tous les intervenants. Pour ce faire, certains processus prévoient des mécanismes d'évaluation et de suivi, et d'autres non. Peut-on les relier à l'évaluation des enseignants? Dans ce cas, existe-t-il des pratiques évaluatives qui permettent de juger de la qualité du travail enseignant? Quels pourraient être les écueils sur lesquels pourrait acher une telle tentative? Quels pourraient être les avantages à instaurer un tel dispositif? Voici le questionnement proposé à l'ébauche de la recherche qui vous sera proposée.

Affiche 8 L'évaluation diagnostique cognitive des compétences en langues – Exemples de quelques modèles de mesures utilisés dans le contexte des tests à grande échelle**Dan Thanh Duong Thi et Nathalie Loye (Université de Montréal)**

L'évaluation diagnostique cognitive vise à fournir aux apprenants un rapport très fin sur leur niveau de maîtrise de diverses habiletés. Cette approche, qui combine la théorie de la cognition et des modèles psychométriques, permet de mesurer et d'interpréter la maîtrise des habiletés évaluées à travers des items reliés à des attributs qui les définissent.

Dans l'évaluation des compétences des langues, l'approche diagnostique cognitive a été appliquée récemment à des tests n'ayant pas pour but le diagnostic, ipar exemple le TOEFL, le TOEFL iBT, etc. Cette application s'appuie sur la construction de matrices Q, qui mettent en relation les attributs à diagnostiquer et les items, dans le but de déterminer la capacité diagnostique de ces tests et ultimement de fournir aux apprenants des rétroactions sur leurs points forts et leurs faiblesses cognitives.

Ainsi, la présente communication vise dans un premier temps à décrire les fonctions et les démarches de l'approche diagnostique cognitive, notamment en lien avec l'évaluation des compétences en langues. Dans un deuxième temps, des exemples d'application de ces modèles de mesure à des tests de langues à grande échelle seront présentés afin d'illustrer la perspective de l'application en pratique de cette approche, ce qui constitue le thème de notre projet de recherche en cours.

Affiche 9 Pertinence des épreuves du MELS et rendement des élèves**France Paquette et Micheline-Joanne Durand (Université de Montréal)**

Les mathématiques et les sciences de la quatrième année de l'école secondaire sont des matières ciblées par la passation d'épreuves communes du MELS pour tous les élèves du Québec. Ces épreuves sont aussi particulières puisqu'elles servent à des fins de certification ou de diplomation et sont corrigées, de façon locale, par chaque enseignant. Depuis l'implantation du nouveau Programme de formation de l'école québécoise, le MELS a proposé des épreuves prototypes, variables d'une année à l'autre et parfois même fort différentes dans leur forme. En juin 2012, les épreuves imposées (uniques)

compteront pour 50 % de la note de l'année pour l'une des compétences disciplinaires. La réussite de ces épreuves devient un enjeu d'une grande importance dans le contexte de la réussite du plus grand nombre. Faut-il se questionner sur la pertinence de ces situations d'évaluation proposées par le MELS et le rendement attendu des élèves?

À partir d'une grille d'analyse descriptive, la pertinence de quatre situations d'évaluation proposées par le MELS en juin 2011 sera appréciée en indiquant les critères d'évaluation retenus, les manifestations observables choisies ainsi que les particularités de la grille. Dans un deuxième temps, une description des situations d'évaluation et une présentation des résultats obtenus par les élèves d'une commission scolaire seront exposées. Finalement, des constats et des propositions concernant les situations d'évaluation proposées par le MELS et les améliorations à apporter à la grille descriptive seront partagés.

Affiche 10 Documenter les pratiques évaluatives des enseignants de la 6^e année du primaire

Mireille **Mouffe**, Caroline **Leblanc** et Micheline-Joanne **Durand** (Université de Montréal)

L'évaluation des compétences a provoqué l'émergence d'une multitude d'approches évaluatives chez les enseignants. Cette situation peut se révéler problématique puisqu'elle les oblige à adopter une forme d'enseignement et d'évaluation pour lesquels ils n'ont pas été formés (Deniger et Kamansi, 2004) et qui s'éloignent de leurs pratiques actuelles (Simon, Turcotte, Ferne et Forgette-Giroux, 2006). Quelles sont les différentes approches évaluatives utilisées par les enseignants de la 6^e année du primaire au Québec? Comment se sont-ils appropriés la démarche d'évaluation des compétences en lecture, écriture et mathématiques et de quelle façon s'y prennent-ils pour porter leur jugement en cours et en fin de cycle? Le but de notre recherche est de documenter les pratiques évaluatives d'enseignants de 6^e année du primaire en regard de l'évaluation des compétences en cours et en fin de cycle à partir de la documentation récente sur le sujet (Scallon, 1998, 2004; Louis, 2004; Durand et Chouinard, 2006). Cette recherche, de type descriptif, nous a permis de recueillir des données à l'aide d'un questionnaire en ligne rempli par 55 enseignants. Par la suite, 14 d'entre eux ont accepté de participer à une entrevue semi-structurée. Les données recueillies ont été traitées afin d'établir des portraits. Ceux-ci nous renseignent sur la perception des enseignants concernant la démarche évaluative, les disciplines scolaires, la différenciation pédagogique, le portfolio, le nouveau bulletin et les cibles de réussite. Les résultats de cette recherche permettent de se faire une idée plus précise sur la diversité des pratiques utilisées à la 6^e année du primaire en fonction, notamment, de l'âge, du genre, de l'expérience et de la formation des enseignants.

Affiche 11 Évaluer le raisonnement clinique avec un test de concordance de script : un exemple en sciences infirmières

Marie-Eve **Latreille** et Éric **Dionne** (Université d'Ottawa)

L'évaluation de construits complexes représente toujours à ce jour un important défi pour les spécialistes en mesure et évaluation. Il est en effet impératif de pouvoir compter sur des instruments valides et fiables permettant de mesurer ces construits complexes. Le test de concordance de script (TCS) est un instrument développé justement pour mesurer l'un de ces construits complexes, à savoir le raisonnement clinique dans le domaine médical. Dans le cadre de notre étude, nous avons développé et expérimenté un TCS visant à rendre compte du raisonnement clinique chez des infirmières, et plus particulièrement dans le domaine de la pédiatrie. Une version en ligne du TCS comportant 15 scénarios et 45 items a d'abord été administrée à des étudiantes de troisième année (n = 30) de formation en sciences infirmières. Les scores obtenus ont été comparés à un groupe d'expertes (n = 15) et à des infirmières autorisées (n = 40). Au cours du processus de validation, nous avons éliminé neuf items, ce qui nous a permis d'améliorer substantiellement la consistance interne de l'instrument bien que des améliorations restent encore à apporter à l'instrument afin d'améliorer sa fiabilité et sa validité. Les résultats démontrent des différences statistiquement significatives à l'égard des scores moyens entre les trois groupes comparés. Nous avons également remarqué des différences notables au regard des trois dimensions mesurées, à savoir la relation thérapeutique, l'intervention et l'évaluation. Nous compléterons notre présentation en exposant les pistes de recherche que nous

comptons mettre en œuvre pour améliorer les propriétés métriques du TCS.

Affiche 12 Analyse de la qualité des épreuves ministérielles de 6^e année de français et de mathématiques de 2010, 2011 et 2012

Marie-Hélène **Asselin** et Micheline-Joanne **Durand** (Université de Montréal)

À des fins de régulation du système et afin de cerner le niveau d'appropriation du programme de formation de l'école québécoise (PFEQ), le ministère de l'Éducation, du Loisir et du Sport (MELS) administre aux élèves en fin de parcours primaire des épreuves uniques depuis 2004. En plus d'aider les enseignants à réguler leur pratique, ces épreuves servent également à réguler le système. Ces épreuves visent à vérifier si les élèves ont développé un niveau de compétence satisfaisant en résolvant des tâches complexes. D'une année à l'autre, les tâches demandées seront différentes, mais par souci d'équité, les épreuves devraient présenter un paramétrage commun et être équivalentes.

Par ailleurs, le jugement professionnel de l'enseignant s'exerce à toutes les étapes de la démarche d'évaluation et s'appuie sur une collecte bien documentée composée de données pertinentes et suffisantes (Lafortune et Allal, 2008; Allal et Mottier Lopez, 2007; Laveault, 2008;). Par contre, la tension lorsque l'enseignant est confronté à des résultats d'épreuves internes différents de ceux d'une épreuve externe est peu documentée.

Cette étude qui vise, dans un premier temps, à analyser les épreuves de français (lecture et écriture) et de mathématiques de la 6^e année du primaire de trois années consécutives a été réalisée afin de déterminer leur qualité à partir de critères relevés chez Roegiers (2003) et Scallon (2004) et à l'aide de tableaux de spécification. Il reste à vérifier si ces épreuves sont à la hauteur de la valeur qu'on leur attribue. Dans un deuxième temps, nous allons aborder la question de la conformité du jugement de l'enseignant à partir d'épreuves internes réalisées tout au long du cycle d'apprentissage avec une épreuve externe réalisée en fin de cycle.

Affiche 13 Indicateurs de qualité de questions écrites : élaboration et premières étapes de validation d'une grille de vérification

Jacinthe **Cloutier**, Yann **Tissier**, Sébastien **Joncas**, André-Sébastien **Aubin** et Christina **St-Onge** (Université de Sherbrooke)

La rédaction de questions d'examens écrits est un défi de taille. Plusieurs balises et consignes peuvent être utilisées pour guider leur élaboration. Néanmoins, il existe peu d'évidences quant à l'impact du respect de ces consignes sur les propriétés psychométriques des questions. Dans une perspective de pratiques appuyées sur des preuves empiriques, il semble nécessaire d'investiguer la relation entre le respect de consignes de rédaction et les propriétés psychométriques des questions.

L'objectif visé est d'élaborer et de valider une grille de vérification pour évaluer la qualité de questions d'examens écrits (respect ou non-respect de consignes de rédaction). Nous avons procédé à une recension des écrits scientifiques qui a permis de relever 108 consignes de rédaction de questions. Un processus itératif a été mis en place pour réviser ces consignes. La grille obtenue est composée de 36 critères et a été appliquée par quatre juges (deux experts de contenu et deux experts de forme) sur 232 questions portant sur des aspects légaux, organisationnels, éthiques et de santé publique en médecine.

Les résultats montrent que les experts de forme étaient plus sévères que les experts de contenu quant à l'évaluation du respect des consignes ($M(ET) = 0,84(0,23)$ et $0,95(0,12)$ respectivement, $t(64,155) = -3,666$, $p = 0,002$). Les taux moyens de pourcentage d'accord observés étaient de 83,8 % (67,8-100 %) pour les experts de forme, 99,1 % (47,1-100 %) pour les experts de contenu et 93,1 % (15,4-100 %) pour l'ensemble des experts. Une analyse d'items (selon la Théorie Classique des Tests) et des analyses de généralisabilité seront aussi réalisées. Ainsi, les résultats révèlent que les deux types d'experts avaient probablement une interprétation différente de certains critères. Qui plus est, ces différences d'interprétation ont certainement diminué le pourcentage d'accord pour lesdits critères. Les travaux à venir consisteront à définir des pistes d'explications et à établir une grille révisée.

Vendredi 16 novembre 2012

ATELIERS 1 et 2 (Vendredi de 9 h à 10 h 30)

Vendredi 9 h	Atelier 1	<p>Démystifier les tests ontariens à grande échelle de l'Office de la qualité et de la responsabilité en éducation</p> <p>Micheline Wilkie (OQRE/Office de la qualité et de la responsabilité en éducation)</p> <p>L'objectif de cet atelier consiste à présenter une courte synthèse des programmes d'évaluation de l'Office de la qualité et de la responsabilité en éducation. L'OQRE est un organisme indépendant du gouvernement provincial, qui évalue le rendement des élèves de l'Ontario en lecture, en écriture et en mathématiques et qui communique les résultats aux parents, aux éducatrices et éducateurs ainsi qu'au gouvernement. Les tests à l'échelle provinciale mesurent les compétences de base à des étapes clés. Les documents d'appui que l'OQRE a publiés ainsi que les processus utilisés pour élaborer les tests et pour les noter seront présentés.</p>
Vendredi 9 h	Atelier 2	<p>Table ronde sur l'évaluation des compétences professionnelles dans une approche programme à la formation des maîtres au Québec</p> <p>Pascale Lefrançois (Université de Montréal), Alain Rousseau (UQTR) et Sylvie Viola (UQAM)</p> <p>Dans le contexte de l'évaluation des programmes effectuée par le Comité d'agrément des programmes de formation à l'enseignement (CAPFE) dans l'ensemble des universités du Québec en 2011, les facultés des sciences de l'éducation ont présenté l'état des lieux du développement des compétences professionnelles dans les différents programmes d'enseignement. Cette table ronde vise à proposer un regard croisé de trois universités relativement aux travaux réalisés pour mettre en place une approche programme qui favorise l'évaluation des 12 compétences professionnelles à l'aide, par exemple, d'un portfolio.</p> <p>Cette table ronde s'articule autour de trois grands thèmes.</p> <p>1- Approche programme</p> <p>Quelles sont les orientations qui guident la mise en place de l'approche programme? Quels changements devraient être apportés à la structure actuelle des programmes pour être en cohérence avec les orientations proposées? Quels sont les principaux acteurs impliqués dans la démarche? Quel mode de fonctionnement est choisi pour concevoir et mettre en place une approche programme?</p> <p>2- Évaluation des compétences professionnelles</p> <p>Comment est assuré le développement de l'ensemble des 12 compétences dans la structure des différents programmes d'enseignement? Quelles modalités d'évaluation sont privilégiées?</p> <p>3- Régulation du processus</p> <p>Quel sont les mécanismes de régulation prévus afin d'apporter rapidement et efficacement les ajustements nécessaires?</p>

BLOC F (Vendredi de 10 h 45 à 11 h 15)

Vendredi 10 h 45	F1	<p>Diagnostiquer les besoins des élèves et choisir le bon outil, un gage de réussite</p> <p>Julie Provencher (MELS)</p> <p>Dans cette communication, nous examinerons comment l'évaluation est étroitement liée à l'enseignement différencié et à l'apprentissage de stratégies. Nous réfléchirons</p>
------------------	-----------	--

sur l'importance de réguler le processus d'apprentissage afin d'améliorer la compétence de l'élève. Nous présenterons les outils utilisés en aide à l'apprentissage dans un contexte de situation de compétence en français, langue seconde au primaire. Nous ferons également état du processus qui a servi de fil conducteur à leur élaboration.

Vendredi 10 h 45 F2 Analyse comparative de la dimensionnalité d'un test de placement en anglais langue étrangère selon la TRI multidimensionnelle et la méthode bifactorielle

Christophe **Chénier** (UQAM)

La dimensionnalité d'un instrument d'évaluation en éducation est un sujet complexe pouvant être traité à l'aide de différentes méthodes d'analyse statistique. La documentation comprend cependant plusieurs résultats contradictoires au sujet de l'efficacité de différentes méthodes d'analyse et de la structure dimensionnelle des tests étudiés (Blais et Laurier, 1995; Immekus et Imbrie, 2008). De plus, la plupart des études existantes s'intéressent essentiellement à la question de l'unidimensionnalité d'un instrument de mesure, sans toujours chercher à approfondir la structure dimensionnelle propre à l'instrument (Froelich et Habing, 2008). Cette présentation exposera l'analyse de la structure dimensionnelle du TCALS, un test de classement en anglais, langue seconde, utilisé afin de classer les nouveaux étudiants d'un cégep dans le cours d'anglais correspondant le mieux à leur niveau de compétence. Cette analyse, faite à l'aide de la théorie de la réponse à l'item multidimensionnelle (TRIM) et de l'analyse bifactorielle, révèle des éléments très intéressants quant à la structure du TCALS lui-même et, au-delà, quant à l'épineux problème de l'unité des compétences de compréhension orale et écrite (Jang et Roussos, 2007). La comparaison des modèles de mesure utilisés est également instructive et permet de mieux saisir les avantages et désavantages des méthodes respectives en ce qui a trait à leur utilisation en éducation.

Vendredi 10 h 45 F3 Oser certifier l'acquisition de compétences par des examens à choix multiple : l'approche de la Commission de la construction du Québec

Jonathan **Bélaïr**, Anne-Andrée **Bourdeau**, Laila **Valin** et Francis **Kelly** (CCQ)

L'industrie de la construction au Québec s'est dotée de différents mécanismes afin d'assurer la compétence des travailleurs. L'apprentissage, prenant différentes formes selon le parcours de chacun, aboutit ultimement à l'évaluation des compétences en vue de la qualification. Or, la certification professionnelle dans un domaine aussi vaste, névralgique et réglementé que celui de l'industrie de la construction représente de nombreux défis. C'est pourtant la réalité quotidienne de l'équipe de l'évaluation de la compétence à la Commission de la construction du Québec (CCQ). En tenant compte de ce contexte unique, l'équipe de l'évaluation de la CCQ certifie l'acquisition de la compétence par des examens à choix multiple. Venez échanger avec les professionnels de la CCQ qui partageront avec vous leur réalité, leur vision ainsi que les choix méthodologiques qui sous-tendent leur approche d'évaluation.

Vendredi 10 h 45 F4 La contribution de l'évaluateur aux processus de production et de crédibilisation du jugement

Marie-Pier **Marchand** (UQAC) et Marthe **Hurteau** (UQAM)

Toute démarche évaluative aspire à générer un jugement (Fitzpatrick, Christie et Mark, 2009) et son succès est étroitement lié à la mise en place d'une démarche rigoureuse ainsi qu'à la performance de l'évaluateur. Dès 1981, dans leur livre *Effective evaluation*, Guba et Lincoln identifiaient l'évaluateur comme une partie prenante importante de la démarche évaluative. Alors que Jabot et Bauchet (2010) associent le succès d'une évaluation aux connaissances, à l'expertise, au jugement ainsi qu'aux habiletés de négociation de l'évaluateur, Perrin (2012) soutient l'influence de son comportement, qui se traduit par la maîtrise de « compétences générales » (*soft skills*).

Ces dernières années, la pression liée à la professionnalisation du domaine de l'évaluation de programmes a donné lieu à la création de divers documents de référence aspirant à établir les orientations et compétences qui balisent la pratique de l'évaluateur, dont les normes du *Joint Committee on Standards for Educational Evaluation*, les principes directeurs de l'*American Evaluation Association* et le référentiel de compétences de la Société Canadienne d'Évaluation en constituant des initiatives reconnues. La communication vise d'abord à présenter le contenu de cette documentation officielle qui guide la pratique de l'évaluateur, puis à créer le rapprochement avec les résultats de l'étude de Hurteau et Houle (2008) qui précisent des qualités de l'évaluateur qui contribuent au bon déroulement de la démarche évaluative relativement à la production du jugement et à sa crédibilisation par les parties prenantes. Les résultats permettent de valider certaines orientations ou composantes déjà relevées dans la documentation et d'avancer d'autres dimensions personnelles à l'évaluateur au caractère plus inédit.

BLOC G (Vendredi de 11 h 30 à 12 h)

- | | |
|------------------|--|
| Vendredi 11 h 30 | <p>G1 Évaluer la langue dans différents pays : une étude comparative d'épreuves ministérielles</p> <p>Pascale Lefrançois (Université de Montréal) et Catherine Brissaud (Université de Grenoble)</p> <p>La plupart des systèmes scolaires se donnent des moyens d'évaluer, à quelques moments charnières du parcours des élèves, la maîtrise de la langue de ces derniers. Pour ce faire, ils proposent des épreuves ministérielles qui jouent, à des degrés divers, un rôle certificatif ou qui permettent de réaliser un bilan des apprentissages.</p> <p>Une recherche documentaire a permis de comparer les épreuves utilisées dans six systèmes scolaires francophones (Québec, Nouveau-Brunswick, Ontario, France, Belgique, Suisse) et trois systèmes anglophones (Québec, Nouveau-Brunswick, Ontario) à trois moments de la scolarité des élèves : la 6^e année du primaire, la 2^e et la 5^e années du secondaire (ou les années qui s'en rapprochaient le plus).</p> <p>L'analyse révèle des différences importantes entre les systèmes scolaires dans les tâches utilisées, qui vont d'une tâche d'écriture hautement contextualisée à des questions d'analyse en dehors de tout contexte de lecture ou d'écriture. De plus, si les critères servant à évaluer la langue sont assez semblables d'un système à l'autre (on y retrouve toujours l'orthographe, la syntaxe, la ponctuation, le vocabulaire), les approches qualitative/globale et quantitative s'opposent dans la manière de porter un jugement sur ces critères.</p> <p>Notre communication s'intéressera donc aux critères d'évaluation de la langue dans les différentes épreuves ministérielles étudiées et à la nature des jugements portés sur ces critères. Une réflexion didactique sur la nature du bilan des apprentissages ainsi effectué et ses conséquences pour les élèves, les enseignants et le système scolaire conclura nos propos.</p> |
| Vendredi 11 h 30 | <p>G3 Une culture scolaire d'évaluation qui comprend les tests provinciaux en Ontario</p> <p>Hélène Boudreau (OQRE/EQAO, Ontario)</p> <p>La politique énoncée par le ministère de l'Éducation de l'Ontario dans le document « Faire croire le succès » décrit les modes d'évaluation : au service de l'apprentissage, en tant qu'apprentissage et de l'apprentissage. Le programme d'évaluation équilibré tient aussi compte des évaluations en salle de classe, à l'échelle de l'école et du conseil scolaire ainsi qu'à grande échelle. De nombreuses initiatives sont en place pour que le personnel scolaire sache incorporer les composantes pédagogiques de l'évaluation</p> |

dans la réalité quotidienne de l'enseignement et de la planification à courte et à longue échéances de l'amélioration de l'apprentissage et du rendement de tous les élèves.

L'Office de la qualité et de la responsabilité en éducation (OQRE) a mis à jour ses rapports informatisés pour faciliter le travail de planification de l'amélioration scolaire de la part de la direction d'école en collaboration avec l'équipe-école et la collectivité scolaire. À partir des trois types de données rapportés aux écoles par l'OQRE, des tableaux et des graphiques affichent et croisent des données démographiques et contextuelles, des données sur l'attitude et le comportement et des données sur le rendement des élèves.

L'exposé permettra la prise de connaissance de la communication améliorée des données et des rapports offerts par l'OQRE aux écoles élémentaires de l'Ontario.

Vendredi 11 h 30

G4

Validité de l'évaluation en apprentissage par cas : comparaison entre l'évaluation automatisée basée sur l'argumentation et l'évaluation manuelle

Geneviève **Gauthier** (Université de l'Alberta)

L'apprentissage utilisant la méthode par cas appliqués a fait ses preuves dans le milieu professionnel. Toutefois, l'évaluation de ces cas représente un défi, car le but de la résolution de cas ne vise pas uniquement à obtenir une bonne réponse. Dans le cadre d'un cours universitaire en pharmacie clinique, nous utilisons une plateforme d'apprentissage interactive pour tester la validité de l'évaluation automatisée basée sur l'argumentation en la comparant à celle d'un enseignant expert.

Une cohorte de 32 étudiants a résolu quatre cas appliqués. Résoudre un cas dans cet environnement interactif requiert de sélectionner des évidences à travers les différents modules permettant de collecter de l'information sur le cas du patient. Afin de vérifier la validité de l'évaluation automatisée, la note résultant de l'évaluation des réponses et de l'ensemble des évidences soutenant ces réponses est comparée à celle donnée par l'enseignant pour chaque performance. La nature de la rétroaction est aussi analysée afin de mieux comprendre les critères et standards de l'enseignant.

Les résultats montrent que les notes automatisées basées sur l'argumentation développée par l'enseignant ne concordent que très minimalement avec la note manuelle donnée par le même enseignant. Les corrélations varient entre 0,2 et 0,01, ce qui suggère que la note automatisée va rarement dans le même sens que celle fournie par l'enseignant. De plus, l'analyse du type de rétroaction donnée par l'enseignant montre une emphase sur le processus de réponse par rapport à la justesse de la réponse et les évidences soutenant ces réponses. Cette étude de cas exploratoire suggère que la validité de l'évaluation automatisée lorsqu'elle est basée sur un modèle d'argumentation n'est pas acceptable, car elle va rarement dans la même direction que l'évaluation manuelle. Les analyses des rétroactions suggèrent que l'accent sur un modèle de processus de résolution pourrait être plus cohérent avec les critères de l'enseignant.

BLOC H (Vendredi de 13 h 45 à 14 h 15)

Vendredi 13 h 45

H1

La méthodologie : le parent pauvre du rapport d'évaluation de programme?

Éric **Dionne** (Université d'Ottawa)

Les résultats d'une opération d'évaluation de programme se communiquent très souvent au moyen d'un rapport qui fait état des diverses étapes suivies par les évaluateurs dans le cadre de leur travail. Ce rapport devrait, en principe, permettre aux détenteurs d'enjeux (*stakeholders*) ainsi qu'aux diverses personnes intéressées par l'opération d'évaluation de suivre le cheminement et les décisions à la fois théoriques et méthodologiques de l'évaluateur et des parties impliquées. En ce qui concerne plus

spécifiquement la section méthodologie du rapport, il s'agit pour l'évaluateur de rendre compte de ses stratégies et de ses choix tout en prenant en considération les contraintes auxquelles il a dû faire face. Les données rendues disponibles concernant la méthodologie peuvent avoir un impact important, notamment, sur la validité apparente du processus d'évaluation. Dans le cadre de nos travaux, nous avons examiné, au moyen d'une analyse de contenu, 25 rapports d'évaluation disponibles publiquement afin de relever les données à caractère méthodologique qui étaient présentées et celles qui étaient omises. Pour ce faire, nous nous sommes appuyés sur un ensemble de propositions afin de bâtir une grille d'analyse que nous avons ensuite utilisée pour examiner les rapports. À titre d'exemple, nous avons regardé les données associées à l'échantillonnage, l'instrumentation, la validité et la fidélité, etc. Les résultats tendent à montrer que les sections dédiées à la méthodologie souffrent de nombreuses lacunes qui sont susceptibles de nuire à la répétabilité des opérations d'évaluation.

Vendredi 13h45

H2 **Analyse d'un dispositif de formation à l'évaluation des apprentissages à la lumière du cadre de la validité globale**

Réjean **Auger** (OgDDASt inc.) et Nicole **Landry** (UQTR)

Dans cette étude prospective, la validité globale (Auger, 2000; Auger et Landry, sous presse) est proposée comme cadre d'analyse privilégié, afin de suivre et améliorer de façon continue un dispositif de formation à l'évaluation des apprentissages. Ce dispositif de formation a comme atout d'être centré sur l'apprenant et de mettre en place des processus et des activités qui mobilisent des capacités de prise de décision et d'action, ainsi que des capacités d'attribution de sens.

Pour concevoir et piloter un dispositif universitaire de formation à l'évaluation des apprentissages, dans une approche par compétences et, qui plus est, pour une évaluation dans des visées plurielles, il apparaît pertinent de recourir à un cadre d'analyse comme celui de la validité globale. Notamment parce que dans sa modélisation du dispositif de formation, un tel cadre permet d'inclure des évidences empiriques issues des neurosciences cognitives : fonctions exécutives (prise de décision et action) et fonctions d'interprétation ou d'attribution de sens (théorie de l'esprit). La TED fait référence notamment à une capacité de réflexion sur ses propres processus de pensée, à savoir la métacognition (Gazzaniga, Ivry et Mangun, 2001).

L'analyse d'un dispositif de formation à la lumière de la validité globale vise à optimiser le potentiel et à baliser les prises de décisions pédagogiques. Un des premiers résultats souhaités est une contribution significative à l'atteinte par le plus grand nombre d'étudiants d'un niveau de compétence de base en évaluation des apprentissages. Le niveau de compétence escompté devrait permettre à ces futurs enseignants de répondre aux exigences actuelles et à venir... d'une évaluation scolaire complexe et en pleine transformation.

Vendredi 13 h 45

H3 **Analyse de la valeur pédagogique et recherche évaluative : évaluation d'outils pédagogiques**

Judith **Beaulieu** et Jacques **Langevin** (Université de Montréal)

Le Groupe DÉFI Accessibilité, en partenariat avec les Éditions du renouveau pédagogique Inc., a travaillé à l'élaboration de manuels scolaires adaptés (FQRSC, 2010). Ces manuels sont à première vue identiques à la version originale : même page couverture, mêmes thèmes et mêmes illustrations aux mêmes pages, mais les textes sont adaptés. Les manuels scolaires adaptés s'adressent à des élèves de troisième année intégrés en classe régulière et qui ont cumulé deux années de retard en lecture. Ces manuels sont offerts en différentes versions : texte simplifié, texte simplifié en orthographe alternatif ou avec ajout de soutien phonétique. L'objectif de cette recherche

est d'évaluer l'efficacité et la productivité des manuels scolaires adaptés durant les six premiers mois de l'essai. La méthodologie de cette recherche est totalement innovatrice, puisqu'elle combine deux méthodes : l'analyse de la valeur pédagogique et la recherche évaluative. L'analyse de la valeur est une méthode rigoureuse de conception et d'évaluation de l'efficacité de produits, procédés ou services. Cette méthode est utilisée dans plusieurs domaines de pointe, dont l'ingénierie et la pharmacie. La recherche évaluative a pour principales visées l'analyse de la pertinence, de la logique, de la productivité, des effets, de la contextualisation et de l'efficacité d'un programme ou d'un produit. Le modèle hybride combinant l'analyse de la valeur pédagogique et la recherche évaluative regroupe les forces liées à la conception propre à l'analyse de la valeur pédagogique (analyse des besoins, analyse fonctionnelle, équipe de conception regroupant des utilisateurs potentiels et cahier des charges fonctionnel) et les forces évaluatives tirées de la recherche évaluative traditionnellement utilisée en éducation (modèle logique opérationnel et analyse d'implantation). Cette communication vise à présenter le modèle méthodologique hybride, l'illustrant par la conception et l'évaluation des manuels scolaires adaptés.

Vendredi 13 h 45

H4

L'analyse de la satisfaction professionnelle chez les nouveaux enseignants et les chevronnés en utilisant la modélisation de Rasch

Carla **Barroso da Costa** et Nathalie **Loye** (Université de Montréal)

Les recherches sur la satisfaction professionnelle datent du début du XX^e siècle. À l'heure actuelle, la satisfaction professionnelle continue à être un des plus importants concepts étudiés à travers des recherches organisationnelles, puisqu'elle est en lien avec l'efficacité des institutions d'enseignement. La satisfaction professionnelle est un état émotionnel positif qui résulte de l'estime personnelle envers le travail ou envers les expériences de travail cumulées. L'expression du plaisir pour les tâches de travail exercées par l'individu peut changer l'atmosphère de travail et, par conséquent, pousser les enseignants à prendre la décision de continuer ou de quitter leur profession.

La présente étude a été développée en utilisant les données des enquêtes pancanadiennes auprès des directions et des enseignants d'écoles primaires et secondaires au Canada, menées en 2006 par une équipe de professeurs de différentes universités du Québec. Une modélisation de Rasch a été appliquée aux données issues du questionnaire destiné aux enseignants. L'objectif est d'analyser l'effet du genre, de la localisation de l'enseignement (Québec ou d'autres provinces au Canada) et du niveau d'enseignement (primaire ou secondaire) sur la satisfaction professionnelle des enseignants d'au plus sept ans d'expérience (N = 711), de ceux qui ont entre sept et 16 ans d'expérience (N = 1 282) et des enseignants avec plus de 16 ans d'expérience dans l'enseignement (N = 2 316).

BLOC I (Vendredi de 14 h 30 à 15 h)

Vendredi 14 h 30

I1

L'examen de l'impact de la position d'un item dans un plan matriciel multiple sur sa difficulté estimée : cas du Programme international pour le suivi des acquis des élèves (PISA)

Koffi-Philippe **Houme** (Conseil des ministres de l'éducation du Canada) et Yunmei **Xu** (OQRE/EQAO, Ontario)

Les évaluations à grande échelle tel que le Programme international de suivi des acquis des élèves (PISA) suscitent un intérêt de plus en plus croissant tant des chercheurs que des responsables politiques. Ces évaluations génèrent des données permettant aux responsables politiques d'examiner, notamment, les programmes d'études. Au

Canada, les provinces utilisent les données de PISA pour appuyer diverses initiatives liées à l'éducation (RHDC, CMEC, et Statistique Canada, 2009).

Les évaluations de PISA sont administrées suivant un plan d'échantillonnage matriciel multiple (Shoemaker, 1973). Cette technique permet de diviser une évaluation en blocs d'items. Ces blocs sont ensuite répartis dans différents cahiers à administrer à différents groupes. La répartition est faite de sorte qu'un même bloc puisse apparaître dans plus d'un cahier et à des positions différentes. Hohensinn et al. (2008) estiment que, lorsque les mêmes items sont présentés suivant différentes séquences à différents groupes, la comparaison des élèves pourrait souffrir du principe de l'équité.

Certains chercheurs ont trouvé que la position d'un item n'affecte pas sa difficulté estimée (Leary et Doran, 1985; Rubin et Mott, 1984; Zwick, 1991). D'autres par contre ont montré que les items placés vers la fin d'un test tendent à présenter une difficulté plus élevée que les autres (Hohensinn et al., 2008; Meyers, Miller et Way, 2009; Walz et al., 2000).

Notre étude vise à évaluer l'effet de la position d'un item sur la difficulté de cet item. En d'autres termes, la difficulté d'un même item varie-t-elle en fonction de sa position dans le cahier?

Nous avons analysé les données de PISA 2009. Les items y ont été regroupés en 13 cahiers de quatre blocs. Chaque bloc d'items apparaît dans quatre cahiers, à la première, seconde, troisième et quatrième positions (OECD, 2012).

La communication vise à présenter nos résultats et à discuter leurs implications.

Vendredi 14 h 30

12 Mesurer et évaluer les compétences et la capacité à innover

Pierre **Gignac** et Nathalie **Loye** (Université de Montréal)

Le développement d'habiletés et de compétences professionnelles en innovation préoccupe de plus en plus les organisations. Dans son rapport « Mesurer l'innovation : un nouveau regard », l'OCDE (2010) insiste sur le besoin d'outils pour prendre en compte l'éventail des compétences en innovation. Aussi, afin de connaître et de comprendre les apprentissages et les actions qui favorisent l'émergence de telles compétences, il devient important de pouvoir les mesurer et les évaluer.

Sur le plan organisationnel, les outils de mesure et d'évaluation des compétences en innovation sont encore peu nombreux (ASTRA, 2007). L'enquête sur l'innovation CIS 2010 de l'OCDE propose modestement « quatre questions environ sur la créativité et les compétences nécessaires à l'innovation (OCDE, 2010) ». En fait, les habiletés et compétences en innovation sont de plus en plus recherchées mais elles demeurent toujours imprécises, chevauchant entrepreneuriat, créativité, gestion, vision stratégique etc.

L'objectif de la présentation est de faire connaître les défis que pose le développement d'un instrument de mesure et d'évaluation des compétences à innover. Plus précisément, la mesure se concentre sur la capacité d'absorption des connaissances des organisations, la capacité de reconnaître la valeur de nouvelles connaissances, externes à l'organisation, de les assimiler et de les exploiter. Concrètement, ceci se traduit par de nouvelles compétences professionnelles, de nouvelles pratiques, de nouveaux services.

Plusieurs habiletés et compétences sous-tendent le concept de capacité d'absorption des connaissances en innovation. Elles seront abordées sous l'angle de la chaîne de connaissances (Holsapple et Jones, 2004) de façon à obtenir une vue d'ensemble des tâches d'évaluation que nécessite le processus d'innovation.

Les élèves d'aujourd'hui doivent apprendre à innover. Les tâches qu'ils auront à exécuter dans l'avenir seront de plus en plus orientées vers l'apprentissage et l'innovation.

Vendredi 14 h 30 **I3** **La formation initiale des maitres et la didactique du français oral : regard sur les grilles d'observation de l'oral**

Annabelle **Caron**, Isabelle **Carignan** et Vincent **Grenon** (Université de Sherbrooke)

Diverses études révèlent que les enseignants ne semblent pas toujours maîtriser les compétences liées à l'enseignement de l'oral (Lebrun, 2005; Maurais, 1999) dans ses différentes dimensions : oral comme objet (Gervais et Mottet, 2007), oral comme médium (Lafontaine et Préfontaine, 2007), oral pragmatique (Maurer, 2001), modèle linguistique à l'oral (Plessis-Bélaïr, 2008), etc. Afin de former de façon plus adéquate les futurs enseignants du primaire à l'oral et d'améliorer, par le fait même, la façon d'enseigner cette compétence auprès des élèves, il nous paraît essentiel de bien cibler la formation initiale des enseignants, d'autant plus que l'oral n'y occupe encore que trop peu place (Lafontaine, 2005). Que cet enseignement-apprentissage soit intégré ou non au programme universitaire, nous estimons que le recours à des instruments d'évaluation validés, ou ayant fait l'objet d'une validation, pourrait constituer un atout pertinent à la formation.

Dans cette communication, nous présenterons une recherche exploratoire faisant appel à une méthode mixte (Tashakkori et Teddlie, 2003) et nous proposerons des éléments descriptifs utiles pour évaluer la qualité du français oral des futurs enseignants au début de leur formation et, ainsi, tenter de déterminer leurs besoins de formation. Dans un premier temps, nous décrirons les indicateurs utilisés dans deux différentes grilles d'observation et d'évaluation de la compétence orale des futurs enseignants. Dans un deuxième temps, nous montrerons les mesures de validité et de fidélité obtenues en évaluant ces deux instruments. Enfin, nous effectuerons un parallèle avec les compétences énoncées dans le référentiel de formation proposé par le ministère de l'Éducation, du Loisir et du Sport du Québec et les besoins de formation des futurs enseignants inscrits en formation initiale des maitres.

Vendredi 14 h 30 **I4** **Le portfolio professionnel : un outil d'évaluation dans la formation des maitres**

Hélène **Meunier**, Gilles **Raïche** et Carole **Raby** (UQAM)

Le portfolio professionnel est implanté à l'Université du Québec à Montréal (UQAM) dans le programme de formation des maitres au baccalauréat en éducation préscolaire et enseignement primaire depuis septembre 2009. Dans ce contexte, le portfolio remplit deux principales fonctions soit une de réflexion permettant de soutenir la pratique réflexive des étudiants ainsi qu'une d'évaluation qui permet la prise en compte du produit et du processus de développement professionnel (Goupil, Petit et Pallascio, 1999). Ainsi, le portfolio devient un outil d'évaluation dans une fonction formative, permettant à l'étudiant de poser un regard critique et de se questionner à la fois sur sa pratique professionnelle et sur le développement de ses compétences (Bélangier, 2009; Gresso et Lomicka, 1999). Il remplit de plus ce rôle d'évaluation dans une visée sommative et d'orientation, visant à illustrer les accomplissements réalisés tout au long de la formation, notamment auprès des formateurs universitaires, des superviseurs de stage et des maitres associés. Finalement, à l'UQAM, il est maintenant associé depuis peu à une visée certificative, essentielle à l'obtention du diplôme. Dans le cadre d'une approche par compétences, le portfolio se distingue alors comme un outil d'évaluation authentique faisant le lien entre les cours et les stages (Bélaïr et Van Nieuwenhoven, 2010).

Dans le contexte de notre projet doctoral, cette communication vise à présenter dans un premier temps le portfolio professionnel comme un outil d'évaluation, et ce, dans les différentes visées propres à l'évaluation des apprentissages. Dans un deuxième temps, nous tenterons d'expliquer en quoi le portfolio permet à la fois d'avoir un portrait du parcours universitaire des étudiants à la formation des maitres ainsi que de

porter un jugement sur le développement de leurs compétences professionnelles.

BLOC J (Vendredi de 15 h 15 à 15 h 45)

- Vendredi 15 h 15** **J1** **Élaborer des grilles d'évaluation à échelles descriptives**
 France **Côté** (Cégep Marie-Victorin) et Hélène **Meunier** (UQAM)
 L'approche par compétences nous incite à évaluer les apprentissages des étudiants à l'aide de situations complexes. Or, peu importe la production requise dans un tel cadre, le processus d'évaluation devient lui aussi un acte complexe. C'est à travers de multiples nuances que les compétences des étudiants se manifestent et c'est pour cette raison que nous avons recours, de plus en plus, à des grilles d'évaluation comportant une dimension qualitative. De par leurs échelles descriptives, ces grilles permettent d'apprécier, dans leur diversité, les productions réalisées par les apprenants. En s'appuyant sur des critères d'évaluation de même que sur des éléments observables, les grilles à échelles descriptives constituent des outils de jugement efficaces qui facilitent l'évaluation et permettent l'attribution d'une note juste et équitable.
 Comment élaborer ces grilles? Quelles sont les balises à respecter? Y a-t-il des pièges à éviter? Quels sont les contextes où les enseignants gagnent à développer de tels outils? La communication vise à partager une synthèse de nos réflexions et de nos pratiques ayant évoluées au cours des dix dernières années. Différents types de grilles seront présentées afin d'illustrer les nuances entre les grilles d'évaluation à échelles descriptives analytiques et les grilles d'évaluation à échelles descriptives globales. Nous discuterons également de la perspective évolutive d'une grille. Une grille de première génération évoluera, à l'usage, vers une grille de deuxième et même de troisième génération. Comment procéder à l'actualisation et à l'amélioration des versions subséquentes de la grille? En terminant, nous ouvrirons la réflexion sur les grilles d'évaluation à échelles descriptives qui sont intégratrices. Nous croyons qu'elles sont porteuses d'avenir.
- Vendredi 15 h 15** **J2** **Étude comparative de la préparation scolaire des enfants à la maternelle en milieux défavorisés montréalais**
 Pierre Lapointe (Université de Montréal)
 Dans une perspective de prévention de l'échec scolaire, les gouvernements provinciaux et fédéral du Canada financent des évaluations, à grande échelle, des populations d'enfants à la maternelle. Ces enquêtes permettent d'illustrer la diversité des modes d'adaptation des enfants à une période critique de leur développement. La préparation scolaire de 10 867 enfants montréalais a été évaluée par des enseignants de la maternelle en 1998 et 2005. L'analyse typologique distingue quatre profils d'enfants : mature, en difficulté d'adaptation, en difficulté d'apprentissage et à risque. Le jugement des enseignants varie selon les caractéristiques personnelles, familiales et scolaires des enfants. Enfin, l'examen comparatif d'évaluations réalisées auprès de deux cohortes d'enfants en milieu socioéconomique défavorisé, à sept années d'intervalle, montre que la répartition des individus, selon ces quatre profils, est relativement stable dans cette population. En conclusion, l'utilité de ce dispositif d'évaluation des enfants à la maternelle est discutée en regard du développement des programmes de prévention en petite enfance.
- Vendredi 15 h 15** **J4** **La conception de tests d'évaluation en lecture, construction de questions et de corrigés**
 Éric **Falardeau**, Pierre **Valois** et Frédéric **Guay** (Université Laval)
 Dans une recherche que nous menons sur l'enseignement et l'évaluation de la lecture

au 1^{er} cycle du secondaire (FQRSC, 2009-2012), nous avons élaboré des grilles d'évaluation formative et sommative avec un groupe d'enseignants de français et de conseillers pédagogiques. Pour vérifier l'effet des outils d'enseignement et d'apprentissage développés sur la compétence en lecture des élèves qui ont suivi pendant une année un enseignement explicite des stratégies de lecture, nous avons construit deux tests de lecture. Ces tests ont été administrés au début et à la fin du protocole d'expérimentation (qui a duré une année) aux 300 élèves du groupe témoin (qui n'ont pas suivi d'enseignement structuré de la lecture) et aux 300 élèves du groupe expérimental.

Dans cette communication, nous présenterons la méthode de travail que nous nous sommes donnée pour élaborer nos deux tests de lecture, qui devaient pouvoir être terminés en 75 minutes. Chaque test devait couvrir toutes les composantes de la compétence déterminées dans nos outils d'enseignement et d'évaluation, présenter un degré de complexité suffisamment discriminant pour distinguer les élèves les plus compétents des moins compétents, et présenter un corrigé détaillé pour assurer une correction uniforme. Nous aborderons différents problèmes auxquels nous avons été confrontés : la conception de questions à choix multiples pour vérifier la compréhension en lecture, l'équilibre entre les questions de compréhension visant les macroprocessus et les microprocessus et la clarté des questions.

Conférence de clôture (Vendredi de 16 h à 16 h 30)

Tendances et visée plurielles en évaluation

Michel Laurier (Université d'Ottawa)

Au cours de cette conférence de clôture, on dégagera les grandes tendances des dernières années en ce qui a trait à l'évaluation en éducation. On abordera le défi de l'évaluation des compétences, l'apport de la technologie en évaluation, les enjeux de l'évaluation à grande échelle, le questionnement des approches quantitatives et la pratique de l'évaluation formative. On verra comment les contributions au cours de la session d'études s'inscrivent dans ces grandes tendances et permettent d'atteindre les « visées plurielles de l'évaluation ».

Michel Laurier est professeur titulaire et doyen à la Faculté d'éducation de l'Université d'Ottawa depuis juillet 2012. Il a été professeur à l'Université de Montréal pendant 20 ans et doyen de la Faculté des sciences de l'éducation de cette université de 2002 à 2010. Il détient un doctorat de l'Université de Toronto et est spécialiste de l'évaluation de la compétence langagière. Il s'intéresse plus particulièrement à l'utilisation des technologies dans l'élaboration de tests de langue. Il est coauteur d'un ouvrage d'introduction à l'évaluation des apprentissages, *Les principes de mesure et de l'évaluation en éducation* et de nombreux articles scientifiques. Il a travaillé à l'élaboration de plusieurs épreuves de langue, de politiques sur l'évaluation des apprentissages et de référentiels de compétence.

Formulaire d'inscription

34^e session d'étude de l'ADMEE-Canada

Les frais d'inscription incluent le dîner du vendredi. **Le dîner du jeudi n'est pas inclus.**

J'assisterai à l'Assemblée générale ADMEE le vendredi 16 novembre. OUI NON

Formulaire électronique : écrire dans la partie ombrée, elle s'agrandira au besoin

Inscription au congrès	<input type="checkbox"/> 300 \$	Participation à la soirée festive	<input type="checkbox"/> 65\$	Total :
Inscription au congrès (étudiant à temps complet)	<input type="checkbox"/> 150 \$	du jeudi, cocher 65\$		

Transmettez-nous au plus tôt votre formulaire d'inscription électronique par courriel et envoyez-le nous par la poste avec votre chèque.

Cochez si adresse : personnelle bureau Date limite de réception des formulaires d'inscription : **9 novembre 2012**

NOM :		PRÉNOM :	
ADRESSE :			
VILLE :		PROVINCE/PAYS :	
CODE POSTAL :		EMPLOYEUR :	
TÉL. PERS. :		OCCUPATION :	
TÉL. BUR. :		TÉLÉCOPIEUR :	
COURRIEL :			

Jeudi 15 novembre 2012

Pour chaque bloc de communication, noircir la case de votre 1^{er} et de votre 2^e choix. (Il sera respecté selon la capacité de la salle.)

Bloc A (13h15-13h45)		Bloc B (14h-14h15)		Bloc C (14h45-15h15)		Bloc D (15h30-16h)		Bloc E (16h15-16h45)	
1 ^{er} choix	2 ^e choix	1 ^{er} choix	2 ^e choix	1 ^{er} choix	2 ^e choix	1 ^{er} choix	2 ^e choix	1 ^{er} choix	2 ^e choix
A1 <input type="checkbox"/>	A1 <input type="checkbox"/>	B1 <input type="checkbox"/>	B1 <input type="checkbox"/>	C1 <input type="checkbox"/>	C1 <input type="checkbox"/>	D1 <input type="checkbox"/>	D1 <input type="checkbox"/>	E1 <input type="checkbox"/>	E1 <input type="checkbox"/>
A2 <input type="checkbox"/>	A2 <input type="checkbox"/>	B2 <input type="checkbox"/>	B2 <input type="checkbox"/>	C2 <input type="checkbox"/>	C2 <input type="checkbox"/>	D2 <input type="checkbox"/>	D2 <input type="checkbox"/>	E2 <input type="checkbox"/>	E2 <input type="checkbox"/>
A3 <input type="checkbox"/>	A3 <input type="checkbox"/>	B3 <input type="checkbox"/>	B3 <input type="checkbox"/>	C3 <input checked="" type="checkbox"/>	C3 <input checked="" type="checkbox"/>	D3 <input type="checkbox"/>	D3 <input type="checkbox"/>	E3 <input type="checkbox"/>	E3 <input type="checkbox"/>
A4 <input type="checkbox"/>	A4 <input type="checkbox"/>	B4 <input type="checkbox"/>	B4 <input type="checkbox"/>	C4 <input type="checkbox"/>	C4 <input type="checkbox"/>	D4 <input type="checkbox"/>	D4 <input type="checkbox"/>	E4 <input type="checkbox"/>	E4 <input type="checkbox"/>

Vendredi 16 novembre 2012

Choisir un des 2 ateliers du vendredi (9h-10h30) : Atelier 1 : Atelier 2 :

Bloc F (10h45-11h15)		Bloc G (11h30-12h)		Bloc H (13h45-14h15)		Bloc I (14h30-15h)		Bloc J (15h15-15h45)	
1 ^{er} choix	2 ^e choix	1 ^{er} choix	2 ^e choix	1 ^{er} choix	2 ^e choix	1 ^{er} choix	2 ^e choix	1 ^{er} choix	2 ^e choix
F1 <input type="checkbox"/>	F1 <input type="checkbox"/>	G1 <input type="checkbox"/>	G1 <input type="checkbox"/>	H1 <input type="checkbox"/>	H1 <input type="checkbox"/>	I1 <input type="checkbox"/>	I1 <input type="checkbox"/>	J1 <input type="checkbox"/>	J1 <input type="checkbox"/>
F2 <input type="checkbox"/>	F2 <input type="checkbox"/>	G2 <input checked="" type="checkbox"/>	G2 <input checked="" type="checkbox"/>	H2 <input type="checkbox"/>	H2 <input type="checkbox"/>	I2 <input type="checkbox"/>	I2 <input type="checkbox"/>	J2 <input type="checkbox"/>	J2 <input type="checkbox"/>
F3 <input type="checkbox"/>	F3 <input type="checkbox"/>	G3 <input type="checkbox"/>	G3 <input type="checkbox"/>	H3 <input type="checkbox"/>	H3 <input type="checkbox"/>	I3 <input checked="" type="checkbox"/>	I3 <input checked="" type="checkbox"/>	J3 <input type="checkbox"/>	J3 <input type="checkbox"/>
F4 <input type="checkbox"/>	F4 <input type="checkbox"/>	G4 <input type="checkbox"/>	G4 <input type="checkbox"/>	H4 <input type="checkbox"/>	H4 <input type="checkbox"/>	I4 <input type="checkbox"/>	I4 <input type="checkbox"/>	J4 <input type="checkbox"/>	J4 <input type="checkbox"/>

Veuillez libeller votre chèque ou mandat-poste au nom d'**ADMEE-Canada** et l'expédier avec le formulaire d'inscription à l'adresse suivante : **ADMEE-Canada**, Département FPE, Bureau 422, Université Laval, 2320, rue des bibliothèques, Québec (Québec) G1V 0A6. ☎ 418- 656-2131 (5053) / 📠 418-656-2885 / ✉ congres@admee.ca

SECTION À REMPLIR OBLIGATOIREMENT SI LES FRAIS D'INSCRIPTION SONT DÉFRAYÉS PAR L'EMPLOYEUR			
L'organisme doit identifier le/la participant/e sur le chèque et le poster avec le formulaire d'inscription.			
Nom de l'organisme :			
Personne-ressource :			
Courriel :		Téléphone :	