



La qualité de l'évaluation

35^e session d'études de l'ADMEE-Canada
Hôtel Palace Royal de Québec

14 et **15** novembre 2013

Le programme est disponible sur le site de l'ADMEE
www.admee.ca



35^e session d'études de L'ADMEE-Canada
14 et 15 novembre 2013
Hôtel Palace Royal de Québec

LA QUALITÉ DE L'ÉVALUATION

C'est avec un plaisir renouvelé que nous vous convions à vous inscrire à la 35^e session d'études de l'Association pour le développement des méthodologies d'évaluation en éducation (ADMEE-Canada), placée sous le thème « *La qualité de l'évaluation* », qui se tiendra à Québec les 14 et 15 novembre prochains.

Si la quête de la qualité constitue le principal enjeu en évaluation, elle n'en demeure pas moins sa raison d'être. En effet, que l'évaluation soit considérée comme un processus ou comme une finalité, qu'elle soit vue comme un moyen ou comme objet, qu'elle soit contextualisée ou non, qu'elle adopte la fonction certificative, formative, diagnostique ou autre, qu'elle s'intéresse aux activités d'un programme, à une institution, au personnel ou aux apprentissages, la quête de sa qualité semble inéluctablement au centre de toute activité humaine.

Qu'est-ce que la qualité ? Existe-t-il des modèles qui traitent de la qualité de l'évaluation ? Comment assurer la qualité de l'évaluation ? Pour qui cette quête de la qualité est-elle indispensable ? Quelles conditions garantissent la qualité de l'évaluation ? L'absence de qualité en évaluation est-elle évidente et comment s'en préserver ? Voilà autant de questions qui méritent qu'on s'y attarde.

Nous osons espérer que ce programme vous offrira l'occasion d'échanges fructueux et enrichissants tout en vous permettant d'explorer de nouveaux horizons de coopération scientifique dans l'univers francophone. Au plaisir de vous rencontrer à Québec les 14 et 15 novembre prochains et de partager avec vous ces moments mémorables.

Je vous souhaite une session d'études mémorable, des échanges scientifiques riches d'enseignement, débouchant sur des perspectives prometteuses.

Pascal Ndinga
Président de l'ADMEE-Canada

HORAIRE DE LA 35^e SESSION D'ÉTUDES

Jeudi 14 novembre 2013

09 h 00	–	10 h 00	Accueil et inscription
10 h 15	–	11 h 45	Atelier 1
11 h 45	–	13 h 15	Dîner libre
13 h 15	–	14 h 15	Communications Blocs A – B
14 h 15	–	14 h 30	Pause
14 h 30	–	15 h 30	Communications Blocs C – D
15 h 30	–	15 h 45	Pause
15 h 45	–	16 h 45	Communications Bloc E – F
16 h 45	–	17 h 00	Pause
17 h 00	–	18 h 00	Communications Blocs G – H
18 h 00	–	18 h 15	Pause
18 h 15	–	19 h 30	Session d'affiches et cocktail

Vendredi 15 novembre 2013

08 h 30	–	09 h 00	Inscription
09 h 00	–	10 h 30	Atelier 2
10 h 30	–	11 h 00	Pause
11 h 00	–	12 h 00	Communications Blocs I – J
12 h 00	–	12 h 15	Pause
12 h 15	–	12 h 45	Communications Bloc K
12 h 45	–	14 h 30	Dîner de l'ADMEE et réunion
14 h 30	–	15 h 00	Clôture et remise des prix

Jeudi 14 novembre 2013

09 h 00 – 10 h 00	Accueil et inscription
10 h 15 – 11 h 45	Atelier 1
11 h 45 – 13 h 15	Dîner libre
13 h 15 – 14 h 15	Communications Blocs A – B
14 h 15 – 14 h 30	Pause
14 h 30 – 15 h 30	Communications Blocs C – D
15 h 30 – 15 h 45	Pause
15 h 45 – 16 h 45	Communications Bloc E – F
16 h 45 – 17 h 00	Pause
17 h 00 – 18 h 00	Communications Blocs G – H
18 h 00 – 18 h 15	Pause
18 h 15 – 19 h 30	Session d'affiches et cocktail

ATELIER 1 10h15-11h45 **Proposition d'un baccalauréat en mesure et évaluation**
 Éric **Dionne** (Université d'Ottawa)

BLOCS A et B (Jeudi de 13 h 15 à 14 h 15)

A1	13h15-13h45 Salle 1	L'évaluation agrégée Martin Lesage , Gilles Raïche et Martin Riopel (UQAM)
A2	13h15-13h45 Salle 2	Le défi de l'évaluation certificative à la fin du secondaire au Brésil : le cas de l'Exame Nacional do Ensino Médio (ENEM) Patrícia Vieira Nunes Gomes (Ministère de l'éducation du Brésil/INEP)
A3	13h15-13h45 Salle 3	Les tests provinciaux en vue de l'amélioration de l'apprentissage de tous les élèves en Ontario Hélène Boudreau (Office de la qualité et la responsabilité en éducation)
A4	13h15-13h45 Salle 4	Penser la qualité de l'évaluation du point de vue des normes socioprofessionnelles des enseignant(e)s Joëlle Morrisette et Serigne Ben Moustapha Diédhiou (Université de Montréal)
B1	13h45-14h15 Salle 1	Les niveaux de complexité des situations évaluatives : fondements de qualité Chantal Roussel (UQAR)
B2	13h45-14h15 Salle 2	Données manquantes et estimations des paramètres d'item selon la théorie classique des tests : une application en sciences de la santé Sébastien Béland (Université de Sherbrooke), Jean-Sébastien Renaud (Université Laval), Gilles Raïche (UQAM), Julie F. Thériault , François Ratté (Université Laval) et Christophe Chénier (UQAM)
B3	13h45-14h15 Salle 3	Modélisation à visée diagnostique des données en lecture des tests d'anglais langue seconde des collèves du Québec Dan Thanh Duong Thi et Nathalie Loye (Université de Montréal)
B4	13h45-14h15 Salle 4	Le portfolio électronique de développement professionnel : un outil indispensable à l'évaluation des compétences professionnelles des étudiants du programme d'enseignement en éducation physique et à la santé (ÉPS) à l'UQTR Alain Rousseau (UQTR)

BLOCS C et D (Jeudi de 14 h 30 à 15 h 30)

C1	14h30-15h Salle 1	La qualité du dossier académique en enseignement supérieur : analyse d'un programme en formation des maîtres Léon Harvey , Marie-Hélène Hébert et Catherine Simard (UQAR)
C2	14h30-15h Salle 2	Caractéristiques d'une bonne pratique d'évaluation en salle de classe : comparaison de quatre cadres de référence Nadine Talbot , Gilles Raïche et Nathalie Michaud (UQAM)
C3	14h30-15h Salle 4	L'évaluation formative du développement de la compétence professionnelle « Agir de façon éthique et responsable dans l'exercice de ses fonctions » chez les stagiaires en enseignement au secondaire Josée-Anne Gouin et Christine Hamel (Université Laval)
C4	14h30-15h Salle 3	« Tu le vois tout de suite qu'il/elle l'a » : l'évaluation des compétences de stagiaires en formation à l'enseignement Martin Massey (Université de Montréal), Pierre Boudreau et Michel Laurier (Université d'Ottawa)
D1	15h-15h30 Salle 4	Les pratiques d'évaluation des apprentissages dans une faculté de génie André-Sébastien Aubin (Université de Sherbrooke)
D2	15h-15h30 Salle 2	Définir des encadrements en évaluation des apprentissages : un pas de plus vers la qualité Mona Ouellet (Commission scolaire du Val-des-Cerfs)
D3	15h-15h30 Salle 1	Les méthodes d'évaluation des apprentissages en enseignement supérieur sont-elles équitables ? Sawsen Lakhal et Serge Sévigny (Université Laval)
D4	15h-15h30 Salle 3	Des outils d'évaluation pour soutenir l'implantation d'une approche par compétences en éducation Pascale Lefrançois et Martin Massey (Université de Montréal)

BLOC E (Jeudi de 15 h 45 à 16 h 15)

E1	15h45-16h15 Salle 1	Examens assistés par ordinateur (EAO) : le virage technologique et la qualité Gilles Leclerc , Joannah Valma , André Martel , Marie-Claude Marin , France Pérusse et Luc Bernier (Université de Montréal)
E3	15h45-16h15 Salle 2	Utiliser le modèle de Guskey pour évaluer l'impact d'une formation en évaluation des apprentissages destinée à des enseignants du primaire et du secondaire : des résultats probants Sylvie Fontaine , Lorraine Savoie-Zajc et Alain Cadieux (UQAO)
E4	15h45-16h15 Salle 3	Analyse des items d'une version francophone adaptée du questionnaire Readiness of health care students for interprofessional learning (RIPLS) Jacinthe Cloutier , Josée Lafrance , Bernard Michallet , Lyson Marcoux et France Cloutier (UQTR)

BLOC F (Jeudi de 16 h 15 à 16 h 45)

F1	16h15-16h45 Salle 3	Vers un meilleur traitement des informations métrologiques Kathye Bélangier (Université de Montréal) et Michel Laurier (Université d'Ottawa)
F2	16h15-16h45 Salle 1	Le concept de validité : évolution et regard critique Nathalie Loye (Université de Montréal), Nathalie André (Université de Poitiers) et Louis Laurencelle (UQAT)
F3	16h15-16h45 Salle 2	La qualité de la formation à l'évaluation des futurs enseignants du primaire et du secondaire au Québec, au Canada et en Europe francophone Pascal Ndinga (UQAM)
F4	16h15-16h45 Salle 4	Analyse des pratiques d'évaluation des apprentissages au Gabon Yvette Moussounda (École Normale Supérieure, Libreville, Gabon)

BLOCS G et H (Jeudi de 17h à 18h)

G1	17h-17h30 Salle 1	Lorsque les nuances linguistiques peuvent affecter la qualité de l'évaluation... comment pensent les évalués ? Roméo Fola et Marielle Simon (Université d'Ottawa), Wolff-Michael Roth (Université Victoria) et Kadriye Ercikan (University of British Columbia)
G2	17h-17h30 Salle 3	Robustesse du modèle de Rasch unidimensionnel à la violation de l'hypothèse d'unidimensionnalité Georges Boade et Jean-Guy Blais (Université de Montréal)
G3	17h-17h30 Salle 2	Modalités et stratégies de correction de productions écrites : analyse d'observations recueillies chez trois enseignants-correcteurs Luc Diarra (Université de Montréal) et Michel Laurier (Université d'Ottawa)
G4	17h-17h30 Salle 4	Comparaison de l'efficacité pédagogique de deux dispositifs d'enseignement du français en contexte béninois Vincent Comlan Gomez et Denis Savard (Université Laval)
H1	17h30-18h Salle 1	Pour des évaluations plus justes dans une perspective probabiliste Martin Riopel (UQAM)
H2	17h30-18h Salle 3	Les présentations par affiches : une analyse à l'aide du modèle Rasch à multifacettes Christophe Chénier et Nadine Talbot (UQAM)
H3	17h30-18h Salle 2	Utilisation de l'interaction sociale aux fins de co-régulation dans le jugement d'évaluation Éric Dionne et Dany Laveault (Université d'Ottawa)
H4	17h30-18h Salle 4	Persévérance scolaire au Québec : la parole aux jeunes et aux intervenants de la région Chaudière-Appalaches Martin Gendron , Julie Melançon , Marie-Hélène Hébert (UQAR), Éric Frenette (Université Laval) et Dominic Simard (UQAR)

Présentation par affiches (Jeudi de 18h15 à 19h30)

-
- Affiche 1 **Enquêtes sur les évaluations des acquis scolaires au Burkina Faso : enjeux et défis méthodologiques**
Dramane **Boly** (Institut supérieur des sciences de la population), Nathalie **Loye** (Université de Montréal)
- Affiche 2 **L'évaluation de la pratique réflexive dans un portfolio électronique chez les stagiaires en enseignement**
Sidi Mohammed **Souidi** et Micheline-Joanne **Durand** (Université de Montréal)
- Affiche 3 **Élaboration et validation d'un outil d'autoévaluation des compétences langagières en écriture chez les immigrants adultes allophones scolarisés en fonction des programmes d'études en français langue seconde**
Xian **Liang** et Micheline-Joanne **Durand** (Université de Montréal)
- Affiche 4 **Documenter les façons de faire des enseignants de la 6e année du primaire dans toutes les étapes de la démarche d'évaluation**
Lakshmee Devi **Ramoo** et Micheline-Joanne **Durand** (Université de Montréal)
- Affiche 5 **L'évaluation des stratégies métacognitives dans l'enseignement et l'apprentissage de l'anglais langue seconde**
Thanh Tu **Nguyen** et Micheline-Joanne **Durand** (Université de Montréal)
- Affiche 6 **Quelles sont les menaces les plus importantes à la qualité des opérations d'évaluation de programme ?**
Maud **Mediell** et Éric **Dionne** (Université d'Ottawa)
- Affiche 7 **L'évaluation des pratiques de différenciation pédagogique des enseignants de la 6e année du primaire favorisant la réussite des élèves**
Wislais **Simervil** et Micheline-Joanne **Durand** (Université de Montréal)
- Affiche 8 **Analyse des contenus essentiels en vue du développement d'un outil visant à baliser l'observation chirurgicale comme activité d'apprentissage**
Isabelle **Raïche**, Éric **Dionne** et Stanley **Hamstra** (Université d'Ottawa)
- Affiche 9 **Comportements d'intimidation et de violence dans le soccer amateur au Québec : la situation des joueurs et des joueuses de 12 à 17 ans inscrits dans un programme sport-études**
Dominic **Simard** et Julien **Marceaux** (UQAR), Éric **Frenette** (Université Laval) et Martin **Gendron** (UQAR)
- Affiche 10 **Impact de la variable "expert" sur la validité des inférences réalisées avec un test de concordance de script**
Jennifer **Smith** et Éric **Dionne** (Université d'Ottawa)

Vendredi 15 novembre 2013

08 h 30 – 09 h 00	Inscription
09 h 00 – 10 h 30	Atelier 2
10 h 30 – 11 h 00	Pause
11 h 00 – 12 h 00	Communications Blocs I – J
12 h 00 – 12 h 15	Pause
12 h 15 – 12 h 45	Communications Bloc K
12 h 45 – 14 h 30	Dîner de l'ADMEE et réunion
14 h 30 – 15 h 00	Clôture et remise des prix

ATELIER 2	9h-10h30	Qualité de l'évaluation à tous les niveaux : développement de synergies nouvelles dans les systèmes éducatifs
	Salle 1	Dany Laveault (Université d'Ottawa)

BLOCS I et J (Vendredi de 11h à 12h)

I1	11h-11h30	Utilisation des grilles d'évaluation critériées pour promouvoir l'apprentissage auto-régulé
	Salle 1	Margarida Romero (Université Laval), Ernesto Panadero (University of Oulu, Finlande)
I2	11h-11h30	Évaluation sommative et qualitative des processus d'apprentissage
	Salle 2	Alexandre Buysse (Université Laval)
I3	11h-11h30	Coût subjectif de l'évaluation : un enjeu pour la qualité ?
	Salle 3	Bernard André (Haute École pédagogique, Suisse)
I4	11h-11h30	L'animation, le trait d'union dans l'évaluation
	Salle 4	Lyne Legault (UQAT)
J1	11h30-12h	La métagrille : un outil pour évaluer la qualité de l'instrumentation pour l'évaluation
	Salle 1	Micheline-Joanne Durand (Université de Montréal)
J2	11h30-12h	La coévaluation des apprentissages : Un état des lieux
	Salle 2	Leah Gauvin (UQAM), Sébastien Béland (Université de Sherbrooke) et Nathalie Michaud (UQAM)
J3	11h30-12h	La qualité de l'évaluation dans l'ingénierie de la formation en Côte d'Ivoire
	Salle 4	Eribrigitte Pale Camara , Ndugumbo Vita , Denis Savard et Jean-Pierre Fournier (Université Laval)
J4	11h30-12h	Bi-référentialisation NCLC/CECRL d'un test : une étude conduite à distance avec la méthode du marque-page
	Salle 3	Sébastien Georges (Centre international d'études pédagogiques)

BLOC K (Vendredi de 12h à 13h)

K1	12h15-12h45 Salle 3	Regard sur les pratiques d'évaluation des enseignantes du préscolaire à partir d'un questionnaire auto-rapporté Monica Boudreau et Marie-Hélène Hébert (UQAR)
K2	12h15-12h45 Salle 2	Élaboration et de validation d'un modèle d'épreuve certificative en lecture et écriture axée sur l'approche par compétences en fin primaire au Sénégal : Problématique Ibrahima Sakho et Micheline-Joanne Durand (Université de Montréal)
K3	12h15-12h45 Salle 1	Analyse factorielle exploratoire d'un questionnaire sur les pratiques d'évaluation en salle de classe Nadine Talbot , Gilles Raïche et Nathalie Michaud (UQAM)
K4	12h15-12h45 Salle 4	Méthodes quantitatives d'estimation des différences entre les hommes et les femmes en sciences sociales Caroline Simard et Claudia Savard (Université Laval)

LES RÉSUMÉS

Atelier | Jeudi 10h15-11h45

Atelier 1

Proposition d'un baccalauréat en mesure et évaluation

Éric **Dionne** (Université d'Ottawa)

La discipline Mesure et Évaluation est en pleine essor depuis de nombreuses décennies. En effet, les connaissances produites par les experts ont augmenté à un rythme soutenu. De nombreuses sous-disciplines ont d'ailleurs émergé. Il suffit de penser, par exemple, à la mesure, à l'évaluation de programme ou encore à l'évaluation des apprentissages. Des spécialistes et des chercheurs œuvrent ou font de la recherche dans de nombreux secteurs comme l'éducation, les sciences sociales ou encore le domaine de la santé. Bref, les connaissances et les contextes de développement de la discipline sont nombreux, complexes et variés. La formation en Mesure et Évaluation s'acquiert, pour l'instant, essentiellement aux études supérieures par le biais des programmes de maîtrise et de doctorat. C'est le cas autant au Canada, aux États-Unis que dans la plupart des pays d'Europe. Cette situation entraîne certains inconvénients comme, notamment, la difficulté des étudiants aux cycles supérieurs à acquérir les concepts de la discipline tout en développant simultanément une posture de chercheur dans cette même discipline. Cette situation peut faire en sorte de créer un « double » standard pour les étudiants qui se lancent dans des études supérieures en mesure et évaluation. Dans le cadre de cette présentation, je présenterai une proposition de baccalauréat de premier cycle en Mesure et Évaluation. Il s'agit d'une proposition visant à développer une meilleure expertise en Mesure et Évaluation tout en préparant adéquatement ces étudiants aux études supérieures. Dans le cadre de ma contribution, je présenterai les fondements de ce programme, sa structure et ses visées.

BLOC A | Jeudi 13h15-13h45

A1

L'évaluation agrégée

Martin **Lesage**, Gilles **Raïche** et Martin **Riopel** (UQAM)

Cette recherche veut approfondir le concept de l'évaluation hiérarchique des apprentissages qui est également dénommée évaluation agrégée. Ce type d'évaluation se définit comme étant un sous-domaine de l'évaluation des équipes où les équipes ont plusieurs niveaux de supervision dans le sens où les chefs d'équipe sont évalués par un ou plusieurs gestionnaires d'équipe selon une hiérarchie arborescente. Le processus d'évaluation permet ainsi d'évaluer différentes compétences qui varient en fonction de la position hiérarchique de l'individu. Bien que ce processus d'évaluation puisse s'exécuter manuellement, son informatisation qui comporte l'enregistrement de la structure arborescente des équipes dans une base de données pourrait accélérer de beaucoup le processus d'évaluation. Afin de remédier à ce problème, une application Internet d'évaluation agrégée qui réside à l'adresse <http://www.labmecas.uqam.ca/cluster/> a été développée selon une méthodologie de recherche développement par le CDAME. Les objectifs de l'actuel projet doctoral sont la mesure des impacts de l'utilisation de l'application « Cluster » au niveau de : (1) l'acquisition des connaissances avec l'application « Cluster » en comparaison avec l'enseignement traditionnel, (2) l'accès aux données d'évaluation sommatives et formatives, (3) la progression des étudiants dans les modules d'enseignement et (4) la mesure de la résistance au changement en ce qui concerne le taux de satisfaction des usagers par rapport à l'utilisation de l'application. Ce projet doctoral voudrait d'abord faire reconnaître le domaine de l'évaluation agrégée par la communauté scientifique pour ensuite développer un logiciel dans le but d'accélérer le processus d'évaluation des organisations et ainsi obtenir une vision plus précise des évaluations de chacun des membres de l'organisation. En effet,

aucune application informatique connue telle que *Moodle*, *WebCT* et *Blackboard* n'est capable d'implanter le processus de l'évaluation agrégée.

A2

Le défi de l'évaluation certificative à la fin du secondaire au Brésil : le cas de l'Exame Nacional do Ensino Médio (ENEM)

Patrícia Vieira Nunes **Gomes** (Ministère de l'Éducation du Brésil/INEP)

Au Brésil, à la fin de l'enseignement secondaire, les élèves passent un examen national créé en 1998 visant à évaluer les apprentissages associés à cette étape de la vie scolaire, aussi bien qu'à contribuer pour la qualité des démarches pédagogiques à partir des résultats obtenus.

À partir de 2009, l'ENEM a commencé à être utilisé comme mécanisme de recrutement pour l'enseignement supérieur dans le but de démocratiser l'accès à l'université. Des changements ont été mis en oeuvre, spécialement en ce qui concerne la correction de l'épreuve de rédaction en langue portugaise pour garantir que le processus à grande échelle soit équitable et efficace.

En 2012, ont été corrigés 4,1 millions de textes avec un groupe de presque 6 000 correcteurs expérimentés, ayant pour base une grille de correction holistique fondée sur la notion de compétences. En ce domaine, le ministère de l'Éducation du Brésil a reformulé le processus de correction des textes de l'épreuve de rédaction en souhaitant garantir la qualité et l'isonomie. Néanmoins, des défis se présentent de façon très forte : quelles propositions théoriques suivre ? Comment élaborer une grille adéquate à la correction à grande échelle ? Quelles balises doivent être employées pour guider l'élaboration des tâches de l'épreuve de rédaction ?

L'effort actuel est celui d'améliorer la grille de correction afin d'assurer la qualité de la correction des textes, tenant en compte les demandes des professeurs, des apprenants et de la société.

A3

Les tests provinciaux en vue de l'amélioration de l'apprentissage de tous les élèves en Ontario

Hélène **Boudreau** (Office de la qualité et de la responsabilité en éducation)

L'Ontario effectue des tests à l'échelle de la province sur les compétences de tous les élèves en lecture, en écriture et en mathématiques à des étapes clés de leur éducation. Les résultats des tests administrés par l'Office de la qualité et de la responsabilité en éducation (OQRE) contribuent à la responsabilisation du système envers le public, et plus que tout, à la planification de l'amélioration continue de l'apprentissage de chaque élève dans le système éducatif financé par les deniers publics provinciaux.

L'OQRE a mis à jour ses rapports informatisés pour faciliter la planification de l'amélioration scolaire faite par les directions d'école en collaboration avec l'équipe-école et la collectivité scolaire. Des tableaux et des graphiques interactifs affichent et croisent des données démographiques et contextuelles, des données sur l'attitude et le comportement des élèves et des données sur le rendement afin de suivre les tendances dans l'apprentissage des élèves et de cerner les domaines du curriculum sur lesquels il faut se pencher d'un cycle d'enseignement à l'autre.

A4

Penser la qualité de l'évaluation du point de vue des normes socioprofessionnelles des enseignants et des enseignantes

Joëlle **Morrisette** et Serigne Ben Moustapha **Diédhiou** (Université de Montréal)

Depuis quelques décennies, sous l'influence de la Nouvelle gestion publique qui s'institutionnalise au Québec comme ailleurs, s'exprime une véritable « fièvre évaluative » (Damon, 2009), soit un renforcement de la fonction politique et administrative de l'évaluation. Celle-ci devient le principal outil de la régulation du système éducatif. Comme le soulève Lessard (2013), cette centralité de l'évaluation s'est accompagnée d'une technicisation du travail évaluatif des enseignants qui subissent une pression à contrôler davantage les apprentissages de leurs élèves. Dans ce contexte, la qualité de l'évaluation repose sur des indicateurs externes, en particulier les résultats des élèves. La recherche

collaborative (Morrissette, 2012) mise sur un postulat d'acteur social compétent (Giddens, 1987), *i.e.* capable d'exercer un contrôle réflexif sur son agir et d'interpréter cet agir, suivant la compréhension qu'il a des circonstances de son action. Elle mise aussi sur l'idée de savoirs partagés entre membres d'une communauté (Morrissette, 2010) qui se coordonnent au quotidien à partir de conventions, *i.e.* de savoirs de leur culture professionnelle, durcis à l'épreuve du temps (Becker, 2004). C'est sur ces bases qu'ont été conduites deux recherches collaboratives (CRSH 2005-2008; FQRSC 2010-2013) s'étant intéressées aux pratiques d'évaluation formative d'enseignantes du primaire. Le dispositif d'enquête, constitué d'un protocole de rétroaction vidéo et d'entretiens de groupe, a permis aux enseignantes de réfléchir à l'efficacité de certaines pratiques, non pas à partir d'une vision techniciste imposée par la nouvelle gestion publique, mais plutôt à partir d'un point de vue « de l'intérieur », dirions-nous, c'est-à-dire de celui du jugement négocié posé à la suite de la coanalyse de situations de pratique (Vinatier, 2010). Aborder la qualité de l'évaluation sous cet angle conduit à mieux comprendre leurs normes professionnelles en matière d'évaluation formative.

BLOC B | Jeudi 13h45-14h15

B1 **Les niveaux de complexité des situations évaluatives : fondements de qualité**

Chantal Roussel (UQAR)

Actuellement, la cohérence de diverses formations est placée au centre des préoccupations pédagogiques se situant dans le prolongement des compétences professionnelles à développer. En éducation, les approches-programmes élaborées par les équipes de formateurs deviennent les projets porteurs de la qualité tant recherchée. Aussi, les acteurs interpellés par les défis de ces approches doivent réfléchir aux fondements, à la structure, aux maillages articulant la construction des savoirs imaginés dans cette perspective. Or, en lien avec l'organisation des apprentissages scolaires, l'évaluation tient une place centrale pour rendre compte des niveaux de développement de ces apprentissages, réalisés par les apprenants.

La recherche que nous avons menée, en effectuant notre thèse doctorale, nous a conduit à examiner les niveaux de complexité inhérents aux situations évaluatives présentes dans un programme d'études de la formation professionnelle au secondaire. Pour ce faire, nous prenions appui sur des éléments théoriques développés par Roegiers (2007) et Roegiers et al. (2010), qui nous ont permis de développer un cadre d'analyse et de concevoir un cadre d'interprétation des situations évaluatives analysées.

En effet, dans la première année du programme de Santé, assistance et soins infirmiers, des niveaux de complexité touchant l'exploration, l'appropriation, l'application procédurale et l'intégration partielle correspondant à des zones d'évaluation, ont été mis au jour. Prenant compte de ces résultats, nous avons examiné sommairement les situations évaluatives de la seconde année de ce programme en vue d'interroger les niveaux de complexité des situations évaluatives qui s'y présentent et de percevoir le déploiement global de l'évaluation dans ce programme.

Pour rendre compte de la qualité de l'évaluation réalisée en adéquation avec les visées prescrites, la compréhension des niveaux de complexité des situations évaluatives semble un facteur fondamental à considérer.

B2

Données manquantes et estimations des paramètres d'item selon la théorie classique des tests : une application en sciences de la santé

Sébastien **Béland** (Université de Sherbrooke), Jean-Sébastien **Renaud** (Université Laval), Gilles **Raïche** (UQAM), Julie F. **Thériault** et François **Ratté** (Université Laval) et Christophe **Chénier** (UQAM)

Bien que leur nombre puisse varier beaucoup selon le contexte, il n'est pas rare de trouver des données manquantes au sein des items qui composent les échelles de mesure des épreuves d'évaluations ou des questionnaires d'enquête. Cette situation fait en sorte que les résultats statistiques obtenus sont biaisés et peuvent malheureusement mener à des interprétations erronées. Les chercheurs ont proposé de nombreuses approches pour traiter les données manquantes de manière à améliorer l'estimation des paramètres d'items et de sujets (Enders, 2010; Graham, 2012). Ainsi, il est possible de penser à la méthode par imputations multiples ou à des approches plus simples de remplacement des données manquantes par une mesure de tendance centrale. Peu de chercheurs ont toutefois porté un regard sur l'effet des données manquantes et les méthodes de traitement de ces données dans le contexte de l'application de la théorie classique des tests. Certains auteurs ont, tout au plus, traité la question sous l'angle de l'impact de la présence de données manquantes sur le calcul du coefficient alpha de Cronbach (Pichette, Béland et Jolani, 2013).

Dans le cadre de cette étude, nous allons évaluer l'efficacité de la méthode par imputations multiples pour minimiser l'impact des données manquantes sur les estimés des paramètres d'items et de sujets obtenus selon la théorie classique des tests. À cet effet, nous analyserons les réponses de 223 étudiants à 25 questions à réponses choisies. À partir de cette matrice de données complète, nous allons générer différents pourcentages de données manquantes selon deux conditions différentes (MCAR et MAR). Ensuite, nous estimerons les paramètres d'items et de sujets et comparerons les estimés avec ceux obtenus à partir des données complètes.

B3

Modélisation à visée diagnostique des données en lecture des tests d'anglais langue seconde des collèges du Québec

Dan Thanh **Duong Thi** et Nathalie **Loye** (Université de Montréal)

Depuis les trois dernières décennies, l'évaluation diagnostique occupe une place importante dans les recherches visant à l'amélioration de l'enseignement et l'apprentissage, car elle permet de dresser un profil très fin des points forts et des faiblesses des élèves.

À l'heure actuelle, il existe peu d'outils conçus spécifiquement pour des visées diagnostiques. Ainsi, plusieurs recherches en évaluation diagnostique sont effectuées à partir des tests standardisés à grande échelle, mais visent à fournir une information diagnostique détaillée sur le niveau des élèves, autres que leurs scores totaux ou leurs rangs.

Au Québec, depuis 2007, des tests adaptatifs en anglais langue seconde sont administrés dans tous les collèges francophones. Leur objectif est de déterminer le niveau des étudiants sur quatre habiletés : Lire, Écrire, Écouter et Parler en utilisant une échelle unique de 1 à 12. Les résultats obtenus permettent un regard d'ensemble sur leur performance. Toutefois, ils renseignent peu sur leurs points forts et leurs faiblesses, donc sur les stratégies pour les améliorer.

Partant de ces limites, notre étude a pour objectif de modéliser des données en lecture des tests d'anglais langue seconde dans les collèges québécois dans une perspective diagnostique. Ainsi, dans le contexte d'un projet doctoral en cours, cette communication visera, dans un premier temps, à décrire le contexte de l'épreuve d'anglais langue seconde dans les collèges québécois. Elle proposera ensuite un cadre conceptuel et méthodologique résumant les possibilités de modéliser des données des tests à grande échelle. Des analyses préliminaires réalisées avec la base de données de cette épreuve en 2008 et 2009 seront présentées permettant de caractériser le processus et de développer les profils diagnostiques en lecture.

B4**Le portfolio électronique de développement professionnel : un outil indispensable à l'évaluation des compétences professionnelles des étudiants du programme d'enseignement en éducation physique et à la santé (ÉPS) à l'UQTR**Alain **Rousseau** (UQTR)

Depuis l'avènement des nouveaux programmes de formation des maîtres orientés vers la professionnalisation et le développement des compétences professionnelles, les instruments utilisés ont été constamment améliorés pour faire en sorte que l'évaluation des étudiants reflète le plus justement possible leur degré de maîtrise des compétences attendues.

Le portfolio retient l'attention parmi les instruments privilégiés en formation. Scallon (2004, p. 311) associe le portfolio à une « collection de travaux réalisés par un élève dans un but précis ». Le portfolio est aussi « un document dans lequel les acquis de formation d'une personne sont définis, démontrés et articulés en fonction d'un objectif » (Legendre 2005, p. 1059).

Compatible avec une approche par compétences, le portfolio est constitué d'un amalgame de traces écrites témoignant du niveau de maîtrise atteint par les élèves et permet une prise en charge de ces derniers dans leurs apprentissages et une meilleure réflexion personnelle (MELS, 2006).

Le programme en enseignement de l'ÉPS de l'UQTR innove depuis l'automne 2011 justement grâce à l'implantation d'un portfolio électronique de développement professionnel auprès de ses étudiants. L'intention de ce portfolio est d'amener l'étudiant à prendre conscience de l'importance des douze compétences nécessaires à la profession enseignante (MEQ, 2001) et à s'engager de façon soutenue dans un développement professionnel. Au cours de ses quatre années de sa formation universitaire, l'étudiant glissera dans son portfolio, tel un journal personnel, des réflexions et des pièces susceptibles de témoigner d'une amélioration ou d'un engagement par rapport à certains constats et, par le fait même, à la maîtrise des compétences professionnelles attendues chez un futur enseignant.

La présente communication permettra de démontrer la démarche mise de l'avant par le programme à l'UQTR et de jeter un regard nouveau et critique sur le portfolio d'un étudiant.

BLOC C | Jeudi 14h30-15h**C1****La qualité du dossier académique en enseignement supérieur : analyse d'un programme en formation des maîtres**Léon **Harvey**, Marie-Hélène **Hébert** et Catherine **Simard** (UQAR)

Les politiques institutionnelles en enseignement supérieur considèrent que les évaluations rapportées au dossier académique ne doivent pas être le résultat d'un processus arbitraire et aléatoire; elles doivent plutôt être équitables, équivalentes, valides et fiables, et en cohérence avec les objectifs d'un programme de formation. De plus, différents organismes d'accréditation tant en Europe qu'aux États-Unis et au Canada énoncent des critères de qualité vis-à-vis des évaluations et exigent que des preuves empiriques soient produites afin de démontrer que ceux-ci sont respectés. Après avoir esquissé les méthodologies mises en place dans la littérature afin de répondre à de telles exigences, nous présenterons une lecture psychométrique des dossiers académiques d'étudiants ayant complété une formation universitaire en enseignement afin de vérifier la validité et la fiabilité des évaluations réalisées, ainsi que la cohérence avec les visées attendues du programme. L'importance de facteurs tels que la cote de rendement au collégial des étudiants, l'équivalence entre deux campus universitaires et l'équité entre les années de formation, le genre et l'âge des étudiants, sera également investiguée. L'échantillon se compose de 873 finissants du programme d'éducation préscolaire et d'enseignement primaire ayant gradué entre 2006 et 2012 dans l'un ou l'autre des deux campus de l'Université du Québec à Rimouski. Les résultats des analyses factorielles exploratoires et confirmatoires suggèrent que les évaluations ne sont pas le résultat d'un processus arbitraire et

aléatoire. Ainsi, l'ensemble des notes des étudiants aux cours se structure en facteurs qui permettent d'inférer, comme attendu, la présence de groupes de compétences. Cependant, si le facteur principal s'avère fiable, prédictible d'une année à l'autre, équivalent en fonction de l'âge et lié à la cote de rendement au collégial, il persiste des différences liées au campus universitaire et au genre. De plus, la fiabilité des mesures issues de groupes spécifiques de compétences nécessite d'être augmentée.

C2

Caractéristiques d'une bonne pratique d'évaluation en salle de classe : comparaison de quatre cadres de référence

Nadine **Talbot**, Gilles **Raïche** et Nathalie **Michaud** (UQAM)

Depuis l'implantation des réformes aux différents ordres d'enseignement scolaires québécois, certaines caractéristiques de l'évaluation des apprentissages en salle de classe sont énoncées, par différents acteurs, sans véritablement être supportées par un cadre de référence. Notamment, la nécessité du caractère authentique, contextualisé, complexe des modalités et des tâches d'évaluation des apprentissages en salle de classe, sont autant de caractéristiques mentionnées dans un contexte d'évaluation des compétences. À celles-ci s'ajoute, par exemple, la nécessité d'évaluer la transversalité des compétences.

Considérant la diversité des caractéristiques énoncées dans la littérature scientifique et professionnelle, quelques auteurs (Scallon, 2004; Tardif, 2006; Raïche, 2006 et Baartman et collab., 2006) les ont rassemblées en catégories nommées, par certains, des principes, des piliers, des dimensions ou encore des aspects. Ces quatre catégorisations ont permis respectivement l'établissement d'un cadre de référence. En comparant ces derniers, nous pouvons y constater des similitudes ainsi que des contradictions. Certains se démarquent par leur nombre élevé de catégories alors que d'autres se différencient par leur nombre moindre de catégories. Cette comparaison permet de conclure à une potentielle utilisation, ou non, de chacun des cadres de référence, de la part des praticiens et des chercheurs.

Ainsi, notre présentation sera subdivisée en deux parties, soit la présentation de chacun de ces cadres de référence et la comparaison entre ceux-ci. Finalement, un constat global sera énoncé.

C3

L'évaluation formative du développement de la compétence professionnelle « Agir de façon éthique et responsable dans l'exercice de ses fonctions » chez les stagiaires en enseignement au secondaire

Josée-Anne **Gouin** et Christine **Hamel** (Université Laval)

La communication portera sur la douzième compétence « Agir de façon éthique et responsable dans l'exercice de ses fonctions » du référentiel québécois « Les orientations de la formation à l'enseignement, les compétences professionnelles » pour les futurs enseignants au secondaire. Notre choix s'est arrêté sur cette dernière parce qu'elle est difficile à cerner (Jutras et Gohier, 2009) tant pour les formateurs de stagiaires que pour les étudiants et qu'elle présente de nombreux défis quant à son évaluation dans le cadre des stages au secondaire. Nous nous intéresserons d'abord à la compréhension de cette compétence, à son déploiement en contexte professionnel et, finalement, à son évaluation formative.

Les référentiels de compétences offrent des points de repère et sont en quelque sorte des guides. En réalité, ils sont insuffisants à une compréhension éclairée et partagée des compétences. Pour développer une représentation partagée de cette compétence professionnelle, nous aborderons les différentes dimensions qui la caractérisent ainsi que les stratégies à mettre en place pour favoriser et enrichir la pratique réflexive des stagiaires. Ces éléments se retrouvent à l'intérieur d'une trajectoire développementale que nous avons développée et présentée à des stagiaires de deuxième année.

Également, nous proposerons des activités de formation qui mobilisent cette compétence et qui permettent le développement du jugement professionnel des étudiants sur les quatre années du baccalauréat. Pour l'inférer, nous nous intéressons aux visées, aux possibilités et aux outils de

l'évaluation formative ainsi qu'au rôle de médiateur des superviseurs de stage.

C4

« Tu le vois tout de suite qu'il/elle l'a » : l'évaluation des compétences de stagiaires en formation à l'enseignement

Martin **Massey** (Université de Montréal), Pierre **Boudreau** et Michel **Laurier** (Université d'Ottawa)

Cette expression du titre, maintes fois entendue, dans nos échanges avec des enseignants associés et des superviseurs universitaires en formation à l'enseignement, nous interroge. Sous l'angle d'une évaluation "globale" de la performance du stagiaire, elle remet en question, ou à tout le moins, semble "remettre en question" le processus et les modalités d'évaluation des stages définis par les programmes de formation. Qu'y a-t-il derrière cette expression ? Que "possède" le stagiaire *i.e.* "le/la stagiaire l'a" ? S'agit-il d'une évaluation globale valide/fiable ? Ce sont là les questions abordées dans notre recherche. Dans le cadre d'une enquête par questionnaire auprès de superviseurs de stage, la question ouverte suivante leur était adressée : "Souvent des enseignants associés et des superviseurs déclarent verbalement que le stagiaire « Il l'a/elle l'a ». Que « possède » un/une stagiaire selon vous quand un enseignant/superviseur emploie une telle expression pour le décrire ? " Plus de 30 superviseurs universitaires de stage, oeuvrant dans diverses universités du Québec et de l'Ontario ont répondu par écrit à la question. Dans un premier temps, les chercheurs, eux-mêmes des formateurs universitaires, liront ces réponses, individuellement, pour tenter d'en saisir le sens; ensuite ils partageront, dans de nombreux échanges, leur lecture et le sens dégagé. Dans un second temps, chacun de ces échanges, enregistrés, servira à la rédaction de notes et de bilans. En parallèle à cette lecture et à cette analyse des données, et pour enrichir cette dernière, les notions de "connaissance" et de "criticism" mises de l'avant par Eisner (1976) de même que celle d'"artistic evaluation" seront utilisées. Les résultats de cette analyse seront présentés. Jusqu'à présent, notre lecture de "surface" des réponses nous porte à croire que les résultats enrichiront les connaissances sur le processus d'évaluation réellement utilisé dans les stages en formation à l'enseignement.

BLOC D | Jeudi 15h-15h30

D1

Les pratiques d'évaluation des apprentissages dans une faculté de génie

André-Sébastien **Aubin** (Université de Sherbrooke)

Notre communication porte sur les pratiques d'évaluation des apprentissages au premier cycle dans une faculté de génie québécoise. Plus spécifiquement, nous documentons deux aspects différents : quelles sont les pratiques d'évaluation des apprentissages des enseignants (professeurs et chargés de cours) et quelles sont les caractéristiques des dispositifs d'évaluation des apprentissages utilisés ?

Deux collectes de données sont utilisées : (a) un questionnaire exhaustif en ligne distribué à tous les enseignants de la faculté et (b) des entrevues individuelles avec 10 enseignants choisis parmi les répondants au questionnaire. Le questionnaire aborde les pratiques d'évaluation (79 énoncés), l'utilisation de projets longs (17 énoncés), l'utilisation de travaux en équipe d'étudiants (5 énoncés), les perceptions de l'évaluation (35 énoncés) et les politiques et règlements touchant l'évaluation (10 énoncés). Les entrevues, d'une durée d'environ 65 minutes, cherchent à décrire en détail les pratiques d'évaluation d'une activité pédagogique.

L'analyse des résultats dresse un portrait précis de l'évaluation des apprentissages dans le contexte choisi en documentant comment les enseignants réalisent les différentes étapes de la démarche évaluative, soit l'identification du but de l'évaluation et du type d'information à rechercher, la planification de l'évaluation, la collecte d'information, l'interprétation, le jugement de la justesse et de la pertinence de l'information, la prise de décision, la communication des décisions et, finalement, la boucle de rétroaction. Elle permet de comparer les caractéristiques des pratiques à celles proposées par les différentes théories en évaluation, dont le paradigme traditionnel, l'approche authentique et

l'approche par compétences. La qualité des pratiques d'évaluation est examinée à l'aide du concept de validité de preuves proposé par Kane (2006).

La présentation s'appuie sur notre projet de recherche doctorale.

D2

Définir des encadrements en évaluation des apprentissages : un pas de plus vers la qualité

Mona **Ouellet** (Commission scolaire du Val-des-Cerfs)

Une des façons d'améliorer la qualité de l'évaluation réside dans l'établissement d'encadrements précis et pertinents au sein de chacune des écoles primaires et secondaires. La définition de ces encadrements est l'occasion de se doter d'une vision commune de l'évaluation et de clarifier les pratiques devant être mises en place par tous les intervenants dans une perspective d'aide à l'apprentissage et de reconnaissance des compétences.

Au cours de la présentation, les participants auront l'occasion de prendre connaissance des éléments incontournables à prendre en compte et des pièges à éviter dans la rédaction de normes et de modalités d'évaluation par une équipe-école. Les conditions nécessaires à la démarche d'élaboration seront aussi présentées afin que cette démarche ne soit plus perçue comme un exercice théorique obligatoire mais bien comme une occasion de réflexion et de confrontation d'idées qui mène à un engagement collectif.

Dans un souci de cohérence et de rigueur en évaluation, il est nécessaire que les intervenants puissent se référer à des balises claires pour appuyer leurs décisions pédagogiques.

D3

Les méthodes d'évaluation des apprentissages en enseignement supérieur sont-elles équitables ?

Sawsen **Lakhal** et Serge **Sévigny** (Université Laval)

Les étudiants n'approchent pas les apprentissages de la même manière à cause, notamment, de leurs différences individuelles. C'est ainsi que dans plusieurs milieux éducationnels, la pédagogie différenciée (Guay, Legault et Germain, 2006; Perrenoud, 2005) a été mise en place pour tenir compte des spécificités des étudiants. En matière d'évaluation des apprentissages, la performance à une méthode d'évaluation donnée semble aussi dépendre de ces différences individuelles, notamment des caractéristiques personnelles des étudiants (Lakhal, Sévigny et Frenette, en révision). Pour être équitable, cette évaluation des apprentissages devrait s'effectuer dans le respect de ces différences individuelles (Ministère de l'Éducation du Québec, 2003) par la mise en place de l'évaluation différenciée (Lakhal, 2012). À cet égard, le ministère de l'Éducation du Québec (2013) énonce : « Il convient (...) que l'évaluation des apprentissages s'inscrive dans cette logique de différenciation, notamment en ajustant les méthodes d'évaluation des apprentissages » (p. 16). Afin d'être en mesure de le faire, les liens entre les différences individuelles des étudiants et la performance aux méthodes d'évaluation devraient être étudiés. La présente communication s'inscrit dans cette perspective et elle a pour objectif de présenter les résultats montrant le lien entre les caractéristiques personnelles des étudiants et la performance à différentes méthodes d'évaluation. Parmi les caractéristiques personnelles des étudiants, la personnalité selon le modèle des cinq facteurs, l'approche d'apprentissage, l'attrait pour les méthodes d'évaluation, le genre, l'âge et la concentration académique ont été retenus.

D4

Des outils d'évaluation pour soutenir l'implantation d'une approche par compétences en éducation

Pascale **Lefrançois** et Martin **Massey** (Université de Montréal)

L'implantation d'une approche par compétences dans des programmes universitaires de 1^{er} cycle pose bon nombre de défis pédagogiques et organisationnels. Parmi ces défis figurent la construction par l'équipe pédagogique d'une vision commune des compétences à développer et, surtout, l'élaboration et l'utilisation d'outils d'évaluation capables de rendre compte du développement de ces compétences et de certifier l'atteinte du profil de sortie attendu au terme de la formation (CAPFE, 2007).

Les programmes de formation initiale des enseignants au Québec sont précisément confrontés à ces défis. D'une durée de quatre ans, alliant annuellement des cours et des stages, ces programmes doivent développer chez leurs étudiants douze compétences professionnelles libellées par le ministère de l'Éducation (MEQ, 2001). Dans le cadre d'une refonte majeure de ses programmes de 1^{er} cycle, la Faculté des sciences de l'éducation de l'Université de Montréal travaille à l'élaboration d'outils d'évaluation des compétences de ses étudiants qui conviennent tant aux cours qu'aux stages. D'abord, un répertoire de manifestations observables des compétences a été construit dans le but d'opérationnaliser les compétences dans un langage commun, utilisé tant dans les plans de cours que dans les grilles d'observation des stages (Raynauld et al., 2012). Ensuite, un portfolio visant à rendre compte de l'intégration des apprentissages a été conçu pour permettre aux étudiants de s'autoévaluer et aux personnes qui les encadrent de leur donner une rétroaction régulière et ciblée. Ces deux outils font partie d'un système pédagogique intégré actuellement en développement, soit une plateforme informatique destinée à soutenir l'implantation de l'approche-programme (Lefrançois, 2012).

Notre communication présentera la réflexion pédagogique à la base de ces outils et quelques-uns des défis qu'elle pose en matière d'évaluation des apprentissages, notamment en termes de qualité de cette évaluation.

BLOC E | Jeudi 15h45-16h15

E1

Examens assistés par ordinateur (EAO) : le virage technologique et la qualité

Gilles **Leclerc**, Joannah **Valma**, André **Martel**, Marie-Claude **Marin**, France **Pérusse** et Luc **Bernier** (Université de Montréal)

Problématique : L'utilisation des technologies pour l'évaluation des apprentissages en sein des programmes par compétence soulève débats et inquiétudes quant à la pertinence des contextes d'utilisation, à la facilité d'implantation, à la sécurité et la fiabilité des approches ainsi qu'à la valeur pédagogique ajoutée. La présentation propose un regard sur deux années d'intégration des EAO au sein du doctorat professionnel de premier cycle en pharmacie (PharmD).

Objectifs : Identifier les avantages/limites pédagogiques et logistiques à l'usage de la technologie.

Décrire les conditions d'implantation et les procédures de soutien optimales nécessaires à l'atténuation des résistances au déploiement des EAO.

Contribuer à l'amélioration continue de la qualité des évaluations.

Méthodes : L'expérimentation, l'observation, les récits d'événements, l'accompagnement personnalisé, les sondages, les rencontres formelles et informelles, les forums de discussion, la collaboration entre pairs et les discussions en groupe ont permis de recueillir les données et témoignages afin de dessiner un portrait réaliste des avantages, limites et conditions pour une intégration optimale des EAO.

Résultats : En plus de provoquer une révision et une transformation des procédures de préparation, d'administration et de correction des examens, l'intégration des EAO a forcé une remise en question importante des rôles et responsabilités de plusieurs intervenants ainsi que des outils, documentation et usages de l'évaluation au sein du PharmD. Les questions de validité et de fidélité des approches et des types d'items d'évaluation préconisés en contexte professionnel ont été abordées et mènent à une transformation progressive des façons de faire.

Discussion/conclusion : Une prise de conscience sérieuse du rôle pédagogique important de l'évaluation émane de l'intégration des EAO. Elle a mené, notamment, à la création du SESAME (service de soutien à la méthodologie en évaluation).

E3

Utiliser le modèle de Guskey pour évaluer l'impact d'une formation en évaluation des apprentissages destinée à des enseignants du primaire et du secondaire : des résultats probants

Sylvie Fontaine, Lorraine Savoie-Zajc et Alain Cadieux (UQAO)

Les enseignants, tant au Québec qu'à l'étranger, considèrent leur compétence en évaluation des apprentissages lacunaire (Fontaine et al. 2012; Mertler, 2009; Kane et al. 2010). Ce constat est inquiétant à la lumière des recherches qui postulent un lien étroit entre l'apprentissage et la réussite des élèves et les pratiques d'évaluation des enseignants (Hattie & Timperley, 2007; Hattie, 2008). Une fois en poste, les enseignants bénéficient de formation continue en évaluation des apprentissages lors de journées pédagogiques. Cependant, les enseignants se plaignent souvent de perdre leur temps lors de ces journées; ils questionnent les liens entre la réalité vécue en classe et les formations qu'ils reçoivent (Guskey, 2000) et remettent en cause l'authenticité du contexte d'apprentissage de ces formations (Tillema, 2007). Il semble ainsi qu'il y ait un paradoxe entre le sentiment mitigé quant à la compétence en évaluation et la formation continue qui vise justement à rehausser le développement de cette compétence. Il est donc pertinent de questionner cette modalité de formation et d'explorer d'autres avenues pour rehausser le sentiment de compétence en évaluation. Trois communautés d'apprentissage professionnelles (CAP) ont été formées en vue de rehausser la compétence en évaluation des apprentissages de 22 enseignants du primaire et du secondaire. Le Five Critical Levels of Professional Development Evaluation de Guskey (2000) a servi de modèle théorique pour évaluer l'impact des CAP sur le rehaussement de la compétence en évaluation. Ce modèle, qui s'inscrit dans la foulée de l'amélioration de la qualité en évaluation, propose une approche exhaustive qui explore la satisfaction des enseignants, comme c'est le cas pour les autres modèles d'évaluation, tout en ayant le souci de cibler aussi les retombées de la formation reçue par l'enseignant pour l'apprentissage et la réussite des élèves. La communication présentera les résultats obtenus pour la 1^{re} année.

E4

Analyse des items d'une version francophone adaptée du questionnaire Readiness of health care students for interprofessionnal learning (RIPLS)

Jacinthe Cloutier, Josée Lafrance, Bernard Michallet, Lyson Marcoux et France Cloutier (UQTR)

La collaboration interprofessionnelle est essentielle dans les professions de la santé (Brock et al. 2013). Voilà pourquoi des programmes universitaires incluent des cours de formation interprofessionnelle (FIP). Toutefois, il n'existe aucun outil francophone validé permettant de mesurer l'impact d'une telle formation sur l'attitude envers la collaboration interprofessionnelle. L'objectif consiste à effectuer une analyse des 19 items d'une version francophone adaptée du RIPLS (Parsell et Bligh, 1999).

Les données ont été recueillies en mars 2013 auprès d'étudiants en sciences de la santé de l'Université du Québec à Trois-Rivières (UQTR) à la suite d'une FIP. Au total, 141 questionnaires papier ont été remplis (taux de réponse = 73 %). Un prétest avait d'abord été effectué pour vérifier la clarté des items (n=20 étudiants en sciences de la santé).

Les alphas de Cronbach obtenus sur les dimensions du questionnaire original présentent une excellente consistance interne : Travail d'équipe et collaboration (neuf items; $\alpha = 0,829$) et Identité professionnelle (huit items; $\alpha = 0,837$). La corrélation de Pearson de la dimension Rôles et responsabilités (deux items) est bonne ($r = 0,372$). La validité de contenu a été vérifiée auprès de trois experts, professeurs en sciences de la santé. Le W de Kendall a été calculé pour vérifier la concordance des juges ($W = 0,575$, $p = 0,028$). L'item 12 de la dimension Identité professionnelle a été rejeté considérant sa faible corrélation item-total.

Selon les résultats, la version francophone adaptée du RIPLS possède de bonnes qualités psychométriques. L'utilisation en est donc recommandée aux universités francophones qui offrent de la FIP. À l'UQTR, ce questionnaire validé sera utilisé pour mesurer l'impact de la FIP (comportant trois cours) sur l'attitude des étudiants en sciences de la santé envers la collaboration interprofessionnelle. Ainsi, une étude longitudinale comportant quatre temps de mesure sera effectuée.

BLOC F | Jeudi 16h15-16h45

F1

Vers un meilleur traitement des informations métrologiques

Kathye **Bélanger** (Université de Montréal) et Michel **Laurier** (Université d'Ottawa)

Au Québec, un test adaptatif informatisé a été élaboré pour attester du niveau des finissants des collèges en français et en anglais langue seconde. Pour les habilités réceptives, les tâches présentées aux candidats sont composées d'un passage à lire ou à écouter, jumelé à un groupe de questions à choix multiple s'y rapportant. Avant d'être intégrées à la banque du test adaptatif, les nouvelles tâches rédigées sont mises à l'essai auprès des collégiens. Les réponses recueillies dans le cadre de cette mise à l'essai sont colligées et font l'objet d'analyses statistiques. L'information métrologique générée est alors interprétée par un expert en mesure et évaluation afin de statuer sur la qualité métrologique des nouvelles tâches et de les situer sur une échelle de compétence en vue de leur intégration à la banque du test.

La recherche a permis de formaliser ce traitement de l'information métrologique effectué par l'expert en mesure et évaluation. À la suite d'entretiens de verbalisation, le raisonnement expert a été modélisé et des règles de raisonnement ont été mises en œuvre dans un algorithme. Cet algorithme a été éprouvé lors du développement des banques de tâches pour les tests adaptatifs en anglais et en français langue seconde. La solution opérationnalisée dans le cadre de cette recherche a ainsi pu être ajustée, raffinée et nuancée au fil de ces mises à l'essai. La formalisation de ce raisonnement expert permet d'uniformiser et d'optimiser le traitement de l'information métrologique, et ce, pour tout test adaptatif informatisé présentant une situation problème à laquelle se greffent plusieurs questions à choix multiples. Concrètement, cette recherche exploratoire à retombées principalement méthodologiques pourra aider à optimiser le développement des banques du test adaptatif. De façon plus large, elle alimente la réflexion sur la démarche d'analyse des tâches dans un contexte certificatif.

F2

Le concept de validité : évolution et regard critique

Nathalie **Loye** (Université de Montréal), Nathalie **André** (Université de Poitiers) et Louis **Laurencelle** (UQAT)

Depuis Binet qui, sans parler de validité, présente une conception pragmatique, utilitaire et empirique de la pertinence d'un test jusqu'à aujourd'hui, le concept de validité a beaucoup évolué. Certains auteurs ont proposé des définitions sans décrire la manière de les opérationnaliser, d'autres réduisent la validité à une pure opération statistique sans appuyer sur un modèle raisonné et explicatif. De plus, certains auteurs, tels Cronbach ou Messick, ont grandement influencé le déploiement du concept de validité. Cette présentation vise à explorer de manière critique quelques facettes du concept de validité et à les mettre en lien avec les démarches proposées pour les explorer en pratique lorsqu'il s'agit d'assurer la qualité de l'évaluation.

F3

La qualité de la formation à l'évaluation des futurs enseignants du primaire et du secondaire au Québec, au Canada et en Europe francophone

Pascal **Ndinga** (UQAM)

La politique d'évaluation découlant du renouveau pédagogique au Québec considère « [...] l'évaluation comme une composante principale du curriculum de la formation générale des jeunes. Cette vision inscrit l'évaluation au cœur des apprentissages. Ainsi, elle constitue un levier pour aider d'une part l'élève à apprendre et d'autre part, l'enseignant à le guider dans sa démarche p. 17. »

On peut se demander cependant si ce rôle pivot de l'évaluation se reflète aussi dans le curriculum de formation des futurs enseignants au primaire et au secondaire. Car, là apparaîtraient les premières manifestations de la qualité potentielle de la formation à l'évaluation des futurs enseignants. Pour

nous en rendre compte, nous avons recensé les programmes de formation des enseignants du primaire et du secondaire des universités du Québec et aussi des universités du Canada et d'Europe francophones, ainsi que les descriptions des cours d'évaluation qui y sont dispensés.

L'analyse des données s'est effectuée à trois niveaux. Dans un premier temps, nous avons quantifié l'offre des cours d'évaluation dans ces programmes de formation des enseignants. Dans un second temps, nous avons analysé les descriptions de ces cours afin d'en estimer la qualité à travers les concepts (contenu) énoncés. À ces deux niveaux d'examen, quantitatif et qualitatif, s'ajoute un troisième niveau, soit la comparaison entre l'offre de cours d'évaluation dans les programmes de formation des enseignants dans les établissements du Canada et d'Europe francophones et celle dans les établissements du Québec.

La présente communication vise à exposer le fruit de cet exercice. Nous discuterons aussi des perspectives éventuelles d'amélioration de la qualité de la formation à l'évaluation des futurs enseignants, notamment par l'amélioration de l'offre des cours en évaluation.

F4

Analyse des pratiques d'évaluation des apprentissages au Gabon

Yvette **Moussounda** (École Normale Supérieure, Libreville, Gabon)

Depuis plus de décennies au Gabon, les résultats scolaires sont unanimement qualifiés de négatifs; plus de 50% d'élèves échouent aux examens nationaux. Certains enseignants (10 pré-enquêtés), n'assurent pas la congruence entre les contenus enseignés et tâches d'évaluation et ne précisent pas toujours les critères aux évalués. Il s'agit en somme d'une absence de valeurs fondamentales et instrumentales en interaction, dont parlent Joanne Durand et Roch Chouinard (2010:26) que sont "l'équité, l'égalité et la justice, la cohérence, la rigueur et la transparence" Comment les enseignants pratiquent-ils l'évaluation des apprentissages? Nous nous voulons analyser des tâches d'évaluation proposées par des enseignants de collège, à la suite de séquences d'enseignement, en vue de cerner leurs pratiques. L'étude sera qualitative.

BLOC G | Jeudi 17h-17h30

G1

Lorsque les nuances linguistiques peuvent affecter la qualité de l'évaluation... comment pensent les évalués ?

Roméo **Fola** et Marielle **Simon** (Université d'Ottawa), Wolff-Michael **Roth** (Université Victoria) et Kadriye **Ercikan** (University of British Columbia)

Bien que les diverses traductions des tests à grande échelle comme ceux du Programme international pour le suivi des acquis (PISA) soient conçues pour être comparables, les nuances linguistiques des populations soumises à ces examens semblent entraver la compréhension des items, et par conséquent, affecter la qualité de l'évaluation. Cette étude fait partie d'un programme de recherche d'envergure plus grande et a pour but d'examiner le degré auquel le français comme langue d'examen constitue un défi pour les élèves francophones en milieu minoritaire. Pour ce faire, nous avons utilisé le protocole de la pensée à voix haute comme une mesure de la qualité des évaluations à grande échelle en termes de comparabilité des résultats entre diverses populations d'élèves francophones en situations minoritaire ou majoritaire, pour tenter de comprendre comment les nuances linguistiques des élèves qui fréquentent les écoles de langue française en Ontario et qui parlent à la maison soit l'une ou l'autre des variantes du français comme le français québécois, ontarien, international, etc., soit une autre langue, peuvent teinter leur compréhension des items de test. La collecte de données a été effectuée dans une école de langue française en Ontario auprès de 35 élèves de 15 ans qui étaient appelés à répondre tout en pensant à voix haute à une quinzaine d'items du PISA préalablement identifiés comme ayant un fonctionnement différentiel et présentant des biais pour ou contre l'une ou l'autre des populations linguistiques. Il s'agissait d'une population très hétérogène dont la langue

d'instruction différerait parfois de celle parlée avec les amis et de celle parlée à la maison. Le protocole exigeait une collecte de données individuelle sous enregistrement audio, qui durait en moyenne 50 minutes. Les résultats préliminaires permettent de constater que la formulation française des items entrave en effet leur compréhension. L'analyse des enregistrements audio nous permettra de comprendre comment.

G2

Robustesse du modèle de Rasch unidimensionnel à la violation de l'hypothèse d'unidimensionnalité

Georges **Boade** et Jean-Guy **Blais** (Université de Montréal)

Le modèle de Rasch est utilisé de nos jours dans de nombreuses applications en sciences sociales et en médecine. Parmi les applications de ce modèle, on trouve l'étude de la qualité psychométrique des items d'un test, le calibrage des items pour les tests adaptatifs, la production des mesures d'habileté en sciences de l'éducation. Il est particulièrement mis à profit dans des enquêtes internationales à grande échelle comme l'enquête PISA (Programme for International Student Assessment). L'une des hypothèses que doivent vérifier l'ensemble des items du test devant mesurer un trait donné est celle de l'unidimensionnalité, c'est-à-dire que tous mis ensemble ne doivent mesurer que le trait en étude, et la réponse que donne un individu à chacun de ces items n'est fonction que du niveau de ce trait chez cet individu. Il se pose donc l'épineuse question de la détermination de la dimensionnalité de l'outil de mesure, car l'objectif étant de ne conserver ensemble que des items concourant à mesurer un seul et même trait. En pratique, les tests auxquels sont soumis les individus ne sont pas strictement unidimensionnels car nos réponses sont aussi conditionnées par nos habitudes et notre milieu. Le plus important selon Stout (1987) est d'avoir un test ayant une dimension dominante, car sinon on devra utiliser des modèles multidimensionnels qui s'avèrent souvent complexes et difficiles à interpréter pour un preneur de décision non expert en mesure. Notre travail a consisté à explorer un ensemble de conditions dans lesquelles le modèle de Rasch unidimensionnel peut produire des mesures acceptables malgré la présence de plusieurs traits déterminants dans les données. Nous avons travaillé avec des données bidimensionnelles simulées, et avons mis à profit le modèle linéaire multiple et les statistiques d'ajustement infit t du modèle de Rasch unidimensionnel.

G3

Modalités et stratégies de correction de productions écrites : analyse d'observations recueillies chez trois enseignants-correcteurs

Luc **Diarra** (Université de Montréal) et Michel **Laurier** (Université d'Ottawa)

Écrire avec l'ordinateur est devenu courant. Certains élèves peuvent avoir développé de meilleures habiletés en modalité de production informatisée (traitement de texte). Par conséquent, pour des raisons d'authenticité et de validité, il devient anachronique de contraindre tous les candidats à la modalité d'évaluation manuscrite. Cependant, prendre en compte les deux modalités d'évaluation manuscrite et informatisée suppose que l'on s'achemine vers la coexistence de plusieurs modalités de correction. Dans ce contexte, une question pertinente est la suivante : le correcteur déploie-t-il les mêmes stratégies suivant qu'il a devant lui une copie manuscrite, imprimée ou affichée à l'écran ?

Le présent travail veut contribuer à l'avancement des connaissances sur la question. Grâce à la verbalisation et une caméra-vidéo, deux enseignants-correcteurs ont été observés en cours de tâche en modalité de correction papier-imprimé, un autre en modalité de correction virtuelle (écran) et tous les trois en modalité de correction papier-manuscrit. Des observations qualitatives ont été recueillies, retranscrites selon leur ordre d'apparition chronologique dans la correction, puis traitées avec le logiciel QDA Miner.

Les résultats de l'analyse des observations qualitatives révèlent que : 1) le correcteur identifie, et les erreurs, et les éléments pertinents positifs, ce qui remet en cause les cadres conceptuels antérieurs focalisés sur la détection de problèmes; 2) la modalité de correction n'a pas d'effet sur les stratégies d'identification d'élément pertinent qui varient plutôt d'un correcteur à l'autre; 3) les stratégies de

rétroaction à l'adresse de l'élève varient par contre sous l'effet des contraintes et des facilités liées à chaque modalité de correction.

G4 **Comparaison de l'efficacité pédagogique de deux dispositifs d'enseignement du français en contexte béninois**

Vincent Comlan **Gomez** et Denis **Savard** (Université Laval)

Le *Just-in-time teaching* apparaît comme un dispositif pédagogique dont l'invention est attribuée à Novak (1999). Il est à la fois un modèle d'enseignement, d'apprentissage et d'évaluation des apprentissages. C'est une pédagogie axée sur la rétroaction entre les travaux hors classe réalisés sur Internet par les apprenants et des activités en classe.

Une adaptation du *Just-in-time teaching* à un contexte caractérisé par une fracture numérique parmi les apprenants donne ce que nous appelons un dispositif d'enseignement Juste-à-temps (DEJAT). Celui-ci est comparé avec le dispositif traditionnellement utilisé.

La présente communication vise à rendre compte de l'effet des dispositifs d'enseignement sur les scores des apprenants, par une analyse de covariance et le calcul du gain d'apprentissage dans une étude de cas.

BLOC H | Jeudi 17h30-18h

H1 **Pour des évaluations plus justes dans une perspective probabiliste**

Martin **Riopel** (UQAM)

L'évaluation telle qu'elle est habituellement pratiquée dans les écoles, les collèges et les universités n'utilise presque jamais les plus récentes méthodes développées par les experts du domaine. Ces méthodes statistiquement optimales, basées sur la théorie de la réponse aux items, sont peut-être trop complexes pour pouvoir être utilisées au quotidien. Pour remédier à ce problème, on peut se demander s'il serait possible de concevoir une méthode de compromis moins complexe et donc plus facilement applicable par les enseignants. Cette nouvelle proposition, à mi-chemin entre les théories classiques et modernes, pourrait permettre d'éliminer des injustices systématiques, souvent ignorées des enseignants, que subissent plusieurs étudiants dans le contexte actuel.

H2 **Les présentations par affiches : une analyse à l'aide du modèle Rasch à multifacettes**

Christophe **Chénier** et Nadine **Talbot** (UQAM)

Plusieurs colloques et congrès académiques sont le théâtre de concours mettant en vedette les présentations orales ou par affiches des présentateurs. Ces concours, où votent normalement les participants du colloque, constituent l'exemple parfait d'une situation d'évaluation pouvant être modélisée à l'aide du modèle Rasch à multifacettes (Linacre, 1989). Ce modèle permet de calculer simultanément le niveau de compétence des candidats, la sévérité des juges et la difficulté des tâches d'évaluation. C'est un modèle très intéressant et largement utilisé par la communauté du *testing*, particulièrement pour les examens de compétences langagières à forts enjeux (Stoynoff, 2012). Or, les évaluations des présentations sont faites par des participants, des juges, ayant potentiellement des niveaux de sévérité différents, et les présentateurs ont eux-mêmes des niveaux de compétence pouvant différer. Il est donc pertinent et intéressant d'appliquer le modèle à ce type de situation. Cette communication présentera les résultats d'une analyse d'une présentation de 41 affiches faite dans le cadre d'un cours universitaire. La présentation a été évaluée par 10 juges, professeurs et étudiants des cycles supérieurs. Les statistiques descriptives pertinentes seront présentées, de même que les indices d'ajustement au modèle et les possibilités d'utilisation de ce modèle à d'autres situations d'évaluation seront évoquées.

H3

Utilisation de l'interaction sociale aux fins de corégulation dans le jugement d'évaluation

Éric Dionne et Dany Laveault (Université d'Ottawa)

L'exercice du jugement, peu importe la forme qu'il peut prendre et le contexte où il peut s'exercer, représente une entreprise intellectuelle complexe et encore aujourd'hui peu ou mal comprise. Dans un premier temps, nous tenterons de définir ce que nous entendons par jugement d'évaluation. Nous dégagerons ensuite les caractéristiques associées aux différents contextes où il s'exerce.

Nous présenterons deux exemples d'outils qui visent à exercer et à étayer le jugement d'évaluation dans un contexte de prise de décision. Le premier a été développé dans le cadre de l'acquisition de compétences en évaluation des apprentissages. Il s'agit d'un outil inspiré des tests de concordance de scripts (TCS) développés dans le domaine médical pour évaluer le jugement clinique. Il a été adapté à la formation de futurs maîtres en évaluation des apprentissages. Les étudiants répondent à différentes questions permettant d'exercer leur jugement d'évaluation en utilisant des télévotants. Quant au second, l'Abaque de Régnier, il s'agit d'un dispositif d'enquête utilisant la technologie web 2.0 permettant la libre expression des points de vue ainsi qu'une rétroaction individuelle et de groupe sur les points de consensus et de dissensus des jugements d'évaluation.

Nous terminerons en montrant comment, dans le cadre d'une procédure favorisant l'interaction sociale, il est possible, avec les outils précédents, de faciliter la corégulation entre juges et d'améliorer la fidélité (concordance interjuges) et la validité du jugement d'évaluation pour la prise de décision.

H4

Persévérance scolaire au Québec : la parole aux jeunes et aux intervenants de la région Chaudière-Appalaches

Martin Gendron, Julie Melançon, Marie-Hélène Hébert (UQAR), Éric Frenette (Université Laval) et Dominic Simard (UQAR)

Le but de la présente étude consiste à recueillir les perceptions de jeunes (préados, ados et jeunes adultes) et d'intervenants quant aux déterminants du système d'éducation québécois pouvant influencer leur persévérance scolaire en vue d'obtenir le diplôme d'études secondaires (ou une équivalence). De fait, il est reconnu que la persévérance scolaire constitue un enjeu sociétal important (Gervais, 2011; GPRSQ, 2009). Le taux de sortie sans diplôme du secondaire entre 2005 et 2007 au Québec était de 30,9 % pour les garçons par rapport à 18,7 % pour les filles (cartodiplome.qc.ca). En Chaudière-Appalaches, ces mêmes taux sont respectivement de 23,5 % et de 11,1 %. Malgré une situation enviable, les acteurs de la région souhaitent comprendre pour faire encore mieux. Une approche qualitative de type focus group est utilisée pour réaliser des entrevues auprès de groupes formés de 5 à 12 personnes. Chacune des quatre commissions scolaires de la région Chaudière-Appalaches a vu les 11 profils de participants différents être rencontrés sur son territoire (n total = 249; n féminin = 124; n masculin = 125). Les canevas d'entrevues semi-dirigées prévus pour 90 à 120 minutes ont été élaborés à partir d'une revue de littérature et validés auprès d'un comité d'experts. À la suite du regroupement et de la catégorisation des éléments de réponse recueillis pour chaque question, les principaux résultats des questions de type « plénière » et « focus group » seront discutés en fonction des enjeux associés à la persévérance scolaire. Notons que parmi les éléments de réponse communs aux divers profils et à toutes les questions, le rôle des enseignants apparaît déterminant au regard de l'opinion des participants. D'autres thèmes figurent dans le discours des participants : la vie sociale, l'intimidation, le réseau de support, la conciliation travail-études, les problèmes de consommation, etc. Enfin, quelques constats et recommandations seront proposés à l'intention des praticiens.

Session d'affiches | Jeudi 18h15

Affiche 1

Enquêtes sur les évaluations des acquis scolaires au Burkina Faso : enjeux et défis méthodologiquesDramane **Boly** (Institut supérieur des sciences de la population), Nathalie **Loye** (Université de Montréal)

Les engagements en faveur de l'« éducation pour tous » se sont traduits au Burkina Faso par la mise en œuvre en 2002 du Plan Décennal de Développement de l'Éducation de Base (PDDEB). Cette politique a eu pour conséquence un accroissement important des effectifs des enfants scolarisés avec une dégradation de la qualité des apprentissages. Pour une meilleure planification en vue d'améliorer la qualité des apprentissages, l'État a mis en place depuis 2004 des enquêtes pour évaluer les acquis scolaires. Ces enquêtes évaluent, conformément aux programmes d'enseignement en vigueur, les habiletés et les compétences des élèves en mathématiques et en français. Les évaluations se font de façon rotative dans les classes du cycle primaire. En 2005, l'évaluation a touché les classes de Cours Préparatoire Première année (CP1) et de Cours Élémentaire Deuxième année (CE2). En 2006, les élèves des classes de Cours Préparatoire Deuxième année (CE2) et Cours Moyen Première année (CM1) ont été évalués et en 2007, les élèves des classes de Cours Élémentaire Première année (CE1) et Cours Moyen Deuxième année (CM2).

Comment se situent ces enquêtes par rapport aux nouvelles méthodes d'évaluation des acquisitions développées dans les grandes enquêtes internationales comme PIRLS, TIMSS et PISA ? Quels sont les enjeux et les défis auxquels ces enquêtes nationales font face surtout en matière de méthodologie ?

La présente communication se propose donc d'apporter des éclairages à ce questionnement.

Affiche 2

L'évaluation de la pratique réflexive dans un portfolio électronique chez les stagiaires en enseignementSidi Mohammed **Souidi** et Micheline-Joanne **Durand** (Université de Montréal)

Au Québec, la pratique réflexive est intégrée dans le référentiel des compétences en formation initiale à l'enseignement. Elle est définie comme une métacompétence pour le développement professionnel (Correa Molina et Gervais, 2012; Perrenoud, 2010; Tardif, 2012). Nous nous intéressons, dans notre recherche, à l'évaluation de la pratique réflexive par l'usage du portfolio électronique en stage d'enseignement, le portfolio se présentant comme un outil d'autoévaluation, d'autoréflexion et de régulation (Scallon, 2004; Lacourse et Hensler, 2008; Tardif, 2006); c'est un outil de soutien pour le développement des compétences professionnelles (Karsenti, Collin et Lepage, 2012; Tardif, 2006). Nos objectifs de recherche visent à montrer la pertinence de l'usage du portfolio électronique en stage d'enseignement pour soutenir la construction de la réflexion chez les stagiaires et l'importance de l'élaboration des grilles d'évaluation descriptives analytiques et globales (Scallon, 2004) pour évaluer la réflexion construite dans un portfolio. Notre démarche méthodologique prévoit une approche qualitative (Anadón, 2006), permettant d'évaluer les traces des réflexions sur le portfolio électronique et les perceptions des sujets sur l'usage de cet outil en stage. Les participants à l'étude seront les stagiaires au primaire et leurs superviseurs. Nous nous inspirons du modèle de Korthagen (2001) pour évaluer les processus réflexifs afin de porter un jugement sur le niveau réflexif de chaque stagiaire en stage.

Affiche 3

Élaboration et validation d'un outil d'autoévaluation des compétences langagières en écriture chez les immigrants adultes allophones scolarisés en fonction des programmes d'études en français langue seconde

Xian **Liang** et Micheline-Joanne **Durand** (Université de Montréal)

Le Québec est une société dont le français est la langue commune dans la vie publique. L'utilisation de la langue française comme moyen de communication touche une grande variété de lieux et de personnes (Pagé, 2005). Les immigrants allophones doivent donc maîtriser adéquatement le français pour pouvoir s'intégrer rapidement à la société et trouver un emploi dans les plus brefs délais (MICC, 2008). Les cours de francisation mettent l'accent sur la communication orale, cependant, les immigrants ne peuvent pas éviter la communication écrite. Nous vivons aujourd'hui dans une société beaucoup plus axée sur l'écrit. Tandis que les rencontres en personne et le téléphone étaient autrefois préconisés, ce sont maintenant les courriels, les messages textes, etc., qui sont davantage utilisés (Thériault, 2011). De plus, même si les immigrants adultes allophones possèdent des diplômes supérieurs provenant de leurs pays d'origine, ils ont des capacités d'écriture nettement limitées en français langue seconde, par exemple pour écrire un chèque, composer des rapports ou des communiqués. La compétence langagière en écriture est directement liée aux capacités cognitives nécessaires à tout apprentissage et à tout fonctionnement professionnel, soit la synthèse d'information, les stratégies d'acquisition de nouvelles connaissances, etc. (Painchaud et Jezak, 1997). Bref, une maîtrise insuffisante de l'écrit, qui peut être l'une des sources de manque de confiance, mène à l'exclusion sociale et professionnelle, autrement dit à l'inégalité sur le marché du travail chez les immigrants allophones (Chicha et Charest, 2008, Vincent, 2004). Notre recherche développement vise à élaborer un outil d'autoévaluation des compétences langagières en écriture afin de faciliter l'apprentissage du français langue seconde chez les immigrants adultes allophones scolarisés et ainsi favoriser leur intégration à la société québécoise.

Affiche 4

Documenter les façons de faire des enseignants de la 6^e année du primaire dans toutes les étapes de la démarche d'évaluation

Lakshmee Devi **Ramoo** et Micheline-Joanne **Durand** (Université de Montréal)

Selon Rey (2008), le changement de curriculum est une « transmutation profonde et quasi anthropologique de la nature des savoirs à transmettre aux jeunes » (p. 7), ayant pour visée la préparation des adultes fonctionnels de la société de demain. Cependant, il y a des questionnements quant à la mise en pratique des changements préconisés dans ce renouveau pédagogique dans les salles de classes québécoises, surtout en ce qui concerne l'évaluation des apprentissages. En effet, les enseignants, qui sont les acteurs de première ligne dans l'implantation des pratiques pédagogiques associées à l'approche par compétences dans les salles de classe, ont sévèrement critiqué son implantation au fil des années (Boutin, 2007).

Selon Fotso (2011), l'adoption d'un paradigme axé sur l'apprentissage a des implications sérieuses en ce qui concerne les façons de faire des enseignants. Dans le cadre de cette étude, nous nous inspirons d'Altet (2002) et de Tardif (1991) pour définir notre compréhension « des façons de faire » des enseignants. En effet, les façons de faire des enseignants seraient des actes, singuliers à chaque enseignant, ayant pour but de recueillir des données pour juger du « développement de l'expertise cognitive » de chaque élève.

Ainsi, le but de cette communication est de faire un portrait des pratiques évaluatives qui ont été mises en place en cours de cycle par les enseignants de la 6^e année du primaire dans un contexte d'approche par compétences. Puis, nous situerons les façons de faire des enseignants dans les cinq étapes de la démarche d'évaluation comme explicitée par Durand et Chouinard (2012). Enfin, nous établirons les ressemblances et les différences entre les pratiques utilisées en mathématiques, en lecture et écriture.

Affiche 5

L'évaluation des stratégies métacognitives dans l'enseignement et l'apprentissage de l'anglais langue seconde

Thanh Tu **Nguyen** et Micheline-Joanne **Durand** (Université de Montréal)

Depuis 1983, plusieurs projets d'enseignement de l'anglais langue seconde ont vu le jour dans le système scolaire québécois. En 2001, le gouvernement québécois imposait l'enseignement de l'anglais à partir de troisième année et, en 2006, en première année dans les écoles primaires françaises. Actuellement, l'anglais langue seconde est enseigné intensivement aux élèves de 6^e année dans plusieurs écoles primaires francophones du Québec, et un projet de loi vise à étendre cette modalité à toutes les écoles à compter de l'année scolaire 2015-2016. Dans le contexte d'une classe d'anglais intensif où 40 % ou plus de temps total d'enseignement est utilisé pour l'enseignement de l'anglais surgit la question du rôle des stratégies. En effet, plusieurs chercheurs en éducation dans ces dernières décennies essaient d'accentuer leurs recherches sur les stratégies d'apprentissage, en particulier les stratégies métacognitives, afin de mieux comprendre comment ces moyens peuvent favoriser l'apprentissage des élèves (Pressley et Afflerbach, 1995; Meijer et al., 2006; Van der Stel et Veenman, 2008). Le terme « métacognition » a été décrit par Flavell (1976) comme « la cognition sur la cognition ». En éducation, la métacognition est considérée comme un élément primordial permettant d'améliorer la capacité de faire face aux situations scolaires chez les apprenants (Pintrich, 1995; Zimmerman et Paulsen, 1995). Les stratégies métacognitives sont définies comme le contrôle et la régulation des connaissances métacognitives des connaissances des tâches et des situations dans lesquelles une tâche complexe est impliquée, pour la gestion des activités mentales (Boulet, 1998). La métacognition s'applique à toutes les activités académiques à l'école. Dans cet esprit, cette recherche vise à comprendre comment les stratégies métacognitives sont évaluées dans l'enseignement et l'apprentissage de l'anglais langue seconde.

Affiche 6

Quelles sont les menaces les plus importantes à la qualité des opérations d'évaluation de programme? Maud **Mediell** et Éric **Dionne** (Université d'Ottawa)

L'évaluation de programme est une discipline grandissante au carrefour de nombreuses disciplines et potentiellement à enjeux critiques. Mener une opération d'évaluation de programme demeure une activité d'une étonnante complexité qui suggère que l'évaluateur ou le groupe d'évaluateurs mette à profit un ensemble considérable et varié de compétences de différents ordres. Stevahn (2005) a d'ailleurs proposé un socle de compétences essentielles divisées en cinq axes principaux : (1) pratique professionnelle, (2) enquête systématique, (3) analyse situationnelle, (4) gestion de projet et (5) pratique réflexive. Ce socle demeure une proposition théorique qui n'a pas été confronté aux données empiriques. Une question demeure, quelles sont les compétences pour lesquelles les évaluateurs se sentent moins outillés pour faire face aux défis quotidiens rencontrés dans le cadre de leur pratique d'évaluateur ? Dans un premier temps, nous proposerons un état de la question sur les difficultés rencontrées par les évaluateurs dans le cadre de leur pratique. Nous tâcherons d'aborder des axes tels que la complexité de la situation réelle d'évaluation, la gestion des contraintes inhérentes au terrain, les interactions avec les parties prenantes, l'utilisation des résultats, le rôle des évaluateurs, etc. Par la suite, nous catégoriserons ces difficultés à la lumière du socle de compétence proposé par la typologie de Stevahn en tâchant d'identifier les axes et les compétences correspondant afin d'élaborer une grille de lecture et ainsi être en mesure d'étudier plus en profondeur et de manière empirique les difficultés rencontrées par les praticiens de l'évaluation, freinant l'élaboration de pratiques évaluatives efficaces, de qualité, et ce, en adéquation avec la réalité de terrain.

Affiche 7

L'évaluation des pratiques de différenciation pédagogique des enseignants de la 6^e année du primaire favorisant la réussite des élèves

Wislais **Simervil** et Micheline-Joanne **Durand** (Université de Montréal)

Depuis plusieurs décennies, le concept de « réussite scolaire » représente une des principales préoccupations de tous les acteurs des systèmes éducatifs notamment avec le renouveau pédagogique. Ce concept est défini comme l'atteinte d'objectifs d'apprentissage liés à la maîtrise des savoirs propres à chaque étape du cheminement scolaire parcouru par l'élève et, ultimement, à l'obtention d'un diplôme ou à l'intégration au marché du travail (Deniger, 2004). Le PRÉE (2010) complète en indiquant que la réussite scolaire est synonyme d'achèvement avec succès d'un parcours scolaire. Les résultats scolaires et l'obtention d'une reconnaissance des acquis (diplôme, certificat, attestation d'études) sont des indicateurs de réussite scolaire. De son côté, Perrenoud (2002) précise que la réussite scolaire est associée à la performance scolaire. Parmi les facteurs qui influencent cette réussite, certains relèvent de l'élève, d'autres de l'enseignant et d'autres encore de l'institution (Roy, 1991). Certains avancent aussi que des pratiques évaluatives telles que celles qui mettent l'emphase sur la différenciation pédagogique visent essentiellement la persévérance et la réussite scolaire (Germain et Proulx, 2007).

Cette recherche vise à examiner en particulier les pratiques de différenciation pédagogique conduites par les enseignants de la 6^e année du primaire qui pourraient favoriser la réussite des élèves spécifiquement en mathématique, en lecture et en écriture.

Par une approche mixte de type exploratoire, nous allons utiliser des données secondaires issues d'une recherche qui a été faite auprès de 55 enseignants de 6^e année du primaire. Les données recueillies à partir du questionnaire seront analysées de manière descriptive et corrélationnelle par le logiciel SPSS. Aussi, les données qualitatives issues des entrevues individuelles seront analysées dans leur contenu.

Affiche 8

Analyse des contenus essentiels en vue du développement d'un outil visant à baliser l'observation chirurgicale comme activité d'apprentissage

Isabelle **Raïche**, Éric **Dionne** et Stanley **Hamstra** (Université d'Ottawa)

Les processus cognitifs des chirurgiens sont reconnus comme étant en grande partie responsables des résultats chirurgicaux. Durant l'acte opératoire, et pour différentes raisons de nature clinique, personnelle ou logistique, le chirurgien enseignant ne les rend pas toujours explicites. Pour les apprentis chirurgiens, reconnaître ces processus cognitifs représente un défi pour lequel ils ne sont généralement pas bien préparés. La salle d'opération étant un milieu riche en information de diverses natures, les apprentis chirurgiens doivent apprendre à fixer leur attention sur les aspects pertinents à la prise de décision. Cette étude vise à définir les indices utilisés par les chirurgiens pour guider leurs décisions intraopératoires. Nous avons choisi de nous concentrer sur les aspects de recueil d'information et de perception visuelle puisqu'il a déjà été démontré que ces habiletés varient entre les experts et les novices. Une analyse de tâches cognitives a été menée auprès de quatre chirurgiens experts d'expériences et de formations différentes. Au total cinq entrevues semi-dirigées ont été conduites avec chaque chirurgien en utilisant un support vidéo durant deux d'entre elles. Les entrevues ont ensuite été analysées de façon qualitative avec la méthode de l'analyse des tâches cognitives proposée par Crandall, Klein et Hoffman (2006). Les chirurgiens experts ont révisé les données brutes et ont confirmé les résultats de l'analyse des données pour procéder à la validation de contenu. Globalement, les chirurgiens utilisent une grande variété d'indices visuels et la réaction des tissus pour guider leur dissection. Des indices cliniques et visuels présents dès le début de la procédure leur permettent généralement de cibler une stratégie particulière à utiliser. En somme, les chirurgiens rencontrés utilisent des processus cognitifs similaires à ceux décrits dans les travaux sur l'expertise.

Affiche 9

Comportements d'intimidation et de violence dans le soccer amateur au Québec : la situation des joueurs et des joueuses de 12 à 17 ans inscrits dans un programme sport-études

Dominic **Simard** et Julien **Marceaux** (UQAR), Éric **Frenette** (Université Laval) et Martin **Gendron** (UQAR)

La présente étude vise à dresser un portrait de la situation de la violence dans le soccer au Québec en fonction du type de violence et du rôle du joueur (témoin, victime, agresseur). Le Questionnaire d'enquête sur le climat dans le soccer amateur au Québec (QECSA; Gendron, Debarbieux, Bodin et Frenette, 2007) a été rempli par 609 joueurs et joueuses U12 à U18 (12 à 17 ans), inscrits à un programme soccer-études. Le QECSA traite de plusieurs variables contextuelles dont les cinq principales sont : (1) types de violence, (2) rôle du joueur face à la violence, (3) lieux d'apparition de la violence, (4) importance du respect des règles du jeu et (5) importance de l'esprit sportif. La présente recherche porte sur les deux premières variables contextuelles soit les types de violence (intimidation verbale (4 items), intimidation physique (4 items), violence physique (3 items)) et le rôle du joueur en tant que témoin de comportements antisportifs. Les résultats seront présentés selon deux approches : la prévalence et la fréquence des trois types de violence (IV, IP, VP). Les résultats concernant la prévalence et la fréquence seront présentés pour l'ensemble des joueurs et selon le genre. Le test t ou t' de Student selon le respect de l'homogénéité des variances ou non sera utilisé afin d'identifier les différences selon le genre. Des analyses comparatives indiquent que les garçons ont tendance à être plus souvent témoins d'IV et d'IP que les filles. Les résultats seront discutés à partir de la littérature existante.

Affiche 10

Impact de la variable « expert » sur la validité des inférences réalisées avec un test de concordance de script

Jennifer **Smith** et Éric **Dionne** (Université d'Ottawa)

Le test de concordance de script (TCS) est un outil développé initialement en médecine visant à rendre compte du jugement clinique des professionnels de la santé. Le TCS est composé d'une vignette présentant une situation authentique complexe et d'une échelle de notation de type Likert sur laquelle les sujets sont invités à exercer leur jugement en y situant leur réponse. La structure des scores de cet outil repose sur l'adéquation entre les réponses des sujets et la distribution des scores d'un panel dit d'experts. En définitive, les résultats visent à mettre en évidence les similitudes existantes entre les processus de pensée (scripts) des sujets à ceux des experts. La popularité des TCS est grandissante et on retrouve des adaptations de ces instruments dans d'autres domaines comme en éducation. Une des variables importantes dans la construction et la validation d'un TCS, peu importe la discipline, est celle associée aux experts. La sélection de ces derniers est une étape névralgique dans la construction d'un TCS puisque la structure des scores dépend de ces derniers. Plusieurs auteurs proposent des balises quant au nombre idéal d'experts à mettre à profit pour la validation, mais il s'agit d'un aspect dont les impacts sont encore mal estimés. Aussi, la notion même « d'expert » peut également poser problème, et ce, surtout dans d'autres domaines que la médecine. Dans le cadre de cette communication, nous ferons le point sur les différentes propositions actuellement connues en lien avec le nombre et la sélection des membres du panel de référence ainsi que l'impact de cette variable « expert » sur la validité des inférences réalisées avec les TCS.

Atelier | Vendredi 9h-10h30

Atelier 2

Qualité de l'évaluation à tous les niveaux : développement de synergies nouvelles dans les systèmes éducatifs

Dany **Laveault** (Université d'Ottawa)

Au cours des vingt dernières années, l'évaluation des apprentissages a pris une importance considérable dans les systèmes d'éducation. Son rôle s'est diversifié et accru dans bien des domaines. Il s'étend de la régulation des apprentissages en salle de classe au pilotage des systèmes d'éducation. Qu'est-ce qui rend légitime ce rôle de plus en plus important de l'évaluation en éducation ? À quels enjeux sont confrontés actuellement les systèmes éducatifs afin que l'évaluation ne serve pas uniquement à rendre des comptes (exigence de redevabilité), mais aussi joue un rôle utile dans l'amélioration et le progrès des apprentissages des élèves (exigence de résultats) ?

Afin de répondre à ces questions, la présentation abordera brièvement chacun des objectifs suivants afin de favoriser le développement d'une perspective globale :

1. Décrire les principaux enjeux de l'évaluation ainsi que leurs incidences sur l'efficacité de l'enseignement et des systèmes éducatifs.
2. Comprendre comment l'évaluation s'articule autour des valeurs d'équité et de responsabilité en éducation.
3. Analyser de manière critique les instruments, méthodes et pratiques d'évaluation en tenant compte des finalités de l'évaluation scolaire.
4. Apprécier les progrès réalisés en matière de qualité de l'évaluation et apporter des solutions à des cas problématiques.
5. Mettre en place des plans d'amélioration qui tiennent compte du contexte institutionnel, des cibles d'amélioration et des données recueillies.

Pour atteindre cet objectif, l'atelier utilisera des exemples tirés des cadres d'évaluation de différentes juridictions nationales et fera référence aux résultats de l'Office de la qualité et de la responsabilité en éducation de l'Ontario ainsi qu'à ceux de l'enquête internationale de l'OCDE (2013) Synergies for Better Learning : An International Perspective on Evaluation and Assessment.

BLOC I | Vendredi 11h-11h30

I1

Utilisation des grilles d'évaluation critériées pour promouvoir l'apprentissage autorégulé

Margarida **Romero** (Université Laval), Ernesto **Panadero** (University of Oulu)

Un des principaux processus d'autorégulation dans la réalisation des activités d'apprentissage est l'autoévaluation. Celle-ci permet de juger la qualité des activités qui sont effectuées et permet de réguler le processus de production et de corriger d'éventuelles erreurs. L'utilisation de rubriques ou grilles d'évaluation critériées permettent à l'apprenant de s'autoévaluer sur la base de critères définis. La grille ou rubrique d'évaluation est un « document subdivisé en critères et en éléments observables, chacun étant accompagné d'une échelle. L'échelle consiste en une succession d'éléments gradués qui correspondent à divers degrés de possession de la qualité visée par le critère » (Côté et Tardif, 2011, p. 8). Les rubriques contribuent à comprendre les critères d'évaluation et réduisent le niveau d'anxiété des apprenants (H. Andrade et Du, 2005). Les rubriques peuvent également contribuer à augmenter le sentiment d'autoefficacité (H. Andrade, Wang, Du, & Akawi, 2009) et permettent améliorer la réalisation des activités d'apprentissage et ses résultats (Alonso-Tapia & Panadero, 2010; H. Andrade, Du, & Mycek, 2010; H. Andrade, Du, & Wang, 2008; Goodrich Andrade, 2001; Hafner & Hafner, 2003) par le biais de l'augmentation des stratégies d'autorégulation (Panadero, Alonso-Tapia, & Reche, 2011). Ainsi la prise de conscience des critères définis dans la rubrique d'évaluation peut permettre à

l'apprenant de développer des stratégies de régulation au cours de la réalisation de l'activité et améliorer ses apprentissages. Malgré ces avantages, fournir aux apprenants une rubrique d'évaluation n'est pas suffisant pour garantir ses effets (Jonsson & Svingby, 2007); les rubriques doivent être expliquées aux apprenants dans le contexte des activités d'apprentissage (Panadero, 2011). Il est donc nécessaire que l'apprenant devienne conscient des objectifs et de l'utilité de l'auto-évaluation pour l'autorégulation et l'apprentissage; il est également nécessaire que l'apprenant apprenne à utiliser les grilles d'évaluation critériées pour améliorer son processus d'autorégulation des apprentissages.

12

Évaluation sommative et qualitative des processus d'apprentissage

Alexandre **Buysse** (Université Laval)

Dans le cadre de la formation scolaire et de la formation professionnelle, la plupart des programmes visent non seulement l'atteinte de compétences disciplinaires mais, globalement, le développement cognitif ou professionnel de l'apprenant. Force est de constater que la plupart des évaluations portent sur les apprentissages et peinent à rendre compte du développement de l'apprenant. Suivant en cela les idées de Vygotski d'évaluer plutôt les processus que les produits (Johnston, 1990), nous nous inspirons de l'évaluation dynamique (Feuerstein, Feuerstein & Falik, 2010). Nous proposons donc une approche de l'évaluation qui ne repose pas sur l'attente de réussite. En effet, dans une perspective développementale, nous proposons de tenir compte du déplacement de la zone proximale de développement (Vygotsky, 1987). Ce déplacement va de pair avec l'intériorisation et la subjectivation des savoirs mais aussi des processus (Buysse & Vanhulle, 2009; Buysse & Renaulaud, 2012). Nous concevons dès lors l'évaluation comme portant à la fois sur les compétences et sur le déplacement de la zone proximale de développement. Cette évaluation devrait permettre de rendre compte à la fois de l'atteinte de certaines compétences et d'un déplacement dans les processus mobilisés par l'apprenant. Nous poserons les enjeux d'une telle problématique et proposerons une tentative d'outil de mesure permettant de rendre compte à la fois de l'atteinte d'objectifs et de révéler des indicateurs de développement scolaire ou professionnel.

13

Coût subjectif de l'évaluation : un enjeu pour la qualité ?

Bernard **André** (Haute École pédagogique)

L'évaluation sollicite les ressources de l'évaluateur à la mesure de la complexité de la situation à évaluer : il s'agit autant de mobiliser ses compétences, principalement cognitives, que du temps, nécessaires à l'analyse de la situation, la planification du recueil de données, l'interprétation de ces données, ainsi que le cas échéant les décisions à prendre. Ce temps nécessaire est sans doute la majeure partie du coût objectif de l'évaluation.

En plus de ces aspects, évaluer mobilise la subjectivité de l'évaluateur, c'est-à-dire ses affects, son histoire, ses valeurs, ses projets, sa perception de lui-même et de son environnement. Cette mobilisation s'accompagne d'un coût subjectif, lié à la nature et à l'intensité de la mobilisation des ressources internes pour faire face aux décisions liées à l'évaluation. S'il est possible, à l'aide de dispositifs expérimentaux et de situations particulières, d'analyser ce coût (Verley, Edouard, et Asselin de Beauville, 2005), notre présentation abordera cette question du coût subjectif de manière qualitative, dans le cadre de l'évaluation des stages faisant partie de la formation des futurs enseignants. Nous mobiliserons deux sources de données : d'une part, nos travaux sur l'investissement subjectif d'enseignants dans leur travail (André, 2013), et, d'autre part, un recueil de données textuelles (bilans de formation et interview) issues de formateurs de terrains en charge de l'évaluation des étudiants en stage dans la formation d'enseignants. L'analyse de ces éléments permettra de faire un lien entre coût subjectif, biais et qualité de l'évaluation dans le cadre de notre recherche.

André, B. (2013). *S'investir dans son travail : les enjeux de l'activité enseignante*. Berne : Peter Lang.

Verley, G., Edouard, J., et Asselin de Beauville, J.-P. (2005). Étude expérimentale du coût subjectif en

théorie bayésienne de la décision. Communication présentée à 12^e rencontre de la Société Francophone de Classification, Montréal.

14

L'animation, le trait d'union dans l'évaluation

Lyne **Legault** (UQAT)

Cette communication repose sur un plan d'animation socioculturelle fondé sur le respect des valeurs du milieu scolaire autochtone dans le but d'évaluer les ressources éducatives des deux écoles d'une communauté autochtone en vue de définir un projet éducatif qui réponde aux exigences du programme gouvernemental, tout en étant spécifique à l'environnement des enfants et adolescents. L'enjeu de l'animation : Évaluer la réussite éducative des élèves, les ressources professionnelles et les installations scolaires en recueillant des données tangibles exprimées par les participants, et ce, avec créativité. Grâce à une démarche d'animation fondée sur l'utilisation d'images, ces derniers ont exprimé leurs visions et leurs perceptions de leur milieu de travail dans un climat d'écoute et de partage. Y a-t-il un parallèle à établir entre le processus de l'animation et celui de la recherche-action ? L'exposé démontrera l'importance de cette démarche d'animation dans un milieu en profonde mutation, ainsi que les résultats de l'évaluation de la vision et de la réussite éducative des écoles d'une communauté autochtone.

BLOC J | Vendredi 11h30-12h

J1

La métagrille : un outil pour évaluer la qualité de l'instrumentation pour l'évaluation

Micheline-Joanne **Durand** (Université de Montréal)

Les changements dans les visées de l'évaluation ont conduit à un questionnement en regard de l'instrumentation utilisée. En effet, dans un contexte d'une évaluation authentique et depuis l'arrivée des programmes axés sur le développement des compétences et l'élaboration de tâches complexes, des outils d'évaluation multiples et variés ont fait leur apparition afin de proposer des solutions aux défis rencontrés par les différents acteurs du milieu de l'éducation : outils de planification, grilles descriptives, portfolio, outil d'autoévaluation et de coévaluation, etc.

Toutefois, en consultant ce qui est proposé par les organismes privés et gouvernementaux et ce qui est utilisé et élaboré dans les différents milieux d'enseignement et de formation, on s'aperçoit que l'instrumentation varie et qu'il est difficile, à partir des exemples donnés, de s'assurer de la qualité de ces outils. On ne peut que constater qu'il y a place pour une étude plus approfondie du sujet. Face à cette situation, une recherche développement a été entreprise afin d'élaborer et de valider des grilles pour analyser la pertinence de ces outils. Des modèles utilisés dans les divers ordres d'enseignement et dans différentes contextes disciplinaires et professionnels ont donc été analysés afin de peaufiner ces grilles. Il fallait par la suite concevoir une grille qui permettait d'analyser et d'évaluer la pertinence de ces premières grilles : une métagrille. Cette grille ne s'adresse pas directement aux productions des élèves mais aux différents outils utilisés pour l'évaluation de ces derniers. Afin d'élaborer cette métagrille, une méthodologie rigoureuse s'avéra nécessaire. C'est la raison pour laquelle l'anasynthèse de Silvern (1972 dans Legendre, 2005) a été choisie. Nous vous présenterons, dans un premier temps, les étapes de la démarche utilisée pour élaborer la métagrille puis, dans un deuxième temps, nous appliquerons cette métagrille à l'analyse de différents outils.

J2

La coévaluation des apprentissages : Un état des lieux

Leah **Gauvin** (UQAM), Sébastien **Béland** (Université de Sherbrooke) et Nathalie **Michaud** (UQAM)

La participation de l'élève à son évaluation s'inscrit dans une logique authentique qui engendre de nombreux avantages (Durand et Chouinard, 2007). Par exemple, elle responsabilise l'élève en plus de lui permettre de développer son esprit d'analyse et sa réflexion. Étrangement, la grande majorité des études portant sur ce thème ont uniquement investiguées le portfolio. Ainsi, il reste beaucoup de travail à faire pour comprendre la portée des autres méthodes telles que l'autoévaluation et la coévaluation.

La communication présentera un résumé de l'éventail des écrits sur la coévaluation. Les auteurs conceptualiseront la coévaluation et présenteront son application dans les écoles primaires, secondaires et les établissements postsecondaires du Québec. La communication démontrera, par les articles analysés ainsi que par l'analyse du contenu des documents de Normes et Modalités d'évaluation des établissements d'enseignement du Québec, la qualité de l'évaluation résultant de cette méthode authentique. De plus, cette communication saura démontrer qu'il y a un besoin urgent de recherche sur la coévaluation, par suite du nombre limité d'écrits sur le sujet comparativement à d'autres méthodes non-traditionnelles d'évaluation des apprentissages.

J3

La qualité de l'évaluation dans l'ingénierie de la formation en Côte d'Ivoire

Eribrigitte **Pale Camara**, Ndugumbo **Vita**, Denis **Savard** et Jean-Pierre **Fournier** (Université Laval)

La qualité de l'évaluation est un aspect essentiel à privilégier dans la réforme du système de formation professionnel en Côte d'Ivoire. L'évaluation fait partie des aspects importants à prendre en compte dans la mise en œuvre des programmes de formation et de développement professionnel. Dans ce sens, une évaluation des besoins de formation (Chagon, Houle et Marcoux, 2009) des acteurs et des bénéficiaires est de mise dans le processus évaluatif.

Au regard des difficultés que rencontre le système de formation professionnelle en Côte d'Ivoire et qui se traduisent par l'inadéquation des contenus de formation, la qualité lacunaire des enseignements et des dispositifs de formation, un certain nombre de questions sont soulevées relativement à l'ingénierie de la formation quant à la qualité du produit et à la pertinence du contenu de la formation. Comment l'ingénierie de la formation participe-t-elle à la formation professionnelle et au développement de secteurs agricoles dans le sens souhaité par la population rurale et urbaine en Côte d'Ivoire ? Le modèle d'ingénierie de la formation en vigueur reste-t-il valide compte tenu de l'évolution de son contexte d'implantation ? Les moyens mis au service de ce programme (matériel didactique, technologies de la communication, outillages) sont-ils efficaces ? Les stratégies didactiques utilisées sont-elles adaptées, compte tenu du contexte de la Côte d'Ivoire et des connaissances les plus récentes ?

Le but de notre exposé consiste à montrer les limites et les forces de l'ingénierie de la formation professionnelle en Côte d'Ivoire et à soulever des questions évaluatives pour améliorer la qualité de l'évaluation de la formation professionnelle.

J4

Bi-référentialisation NCLC/CECRL d'un test : une étude conduite à distance avec la méthode du marque-page

Sébastien **Georges** (Centre international d'études pédagogiques)

Le centre international d'études pédagogiques (CIEP), opérateur public du ministère de l'Éducation nationale française, assure la gestion administrative et la conception pédagogique de certifications en langue française, dont le test de connaissance du français (TCF). Cette communication présente l'étude de mise en correspondance des scores du TCF avec les niveaux de compétence langagière canadiens (NCLC). Originellement, le TCF est adossé aux six niveaux du Cadre européen commun de référence

pour les langues. Conduite par le CIEP, cette étude permet à Citoyenneté Immigration Canada de disposer, pour les candidats à l'immigration au Canada, d'un outil de mesure des compétences langagières valide, fidèle, sensible et équitable, directement adossé aux NCLC. Dix-neuf experts des NCLC et du français ont participé à cette étude. Informés du contexte, de l'enjeu et de la méthodologie, leur participation s'est intégralement déroulée à distance. Chaque groupe d'experts (3) a reçu un livret comprenant soit 36 items de compréhension orale (accompagnés des bandes sons correspondantes), soit 31 items de maîtrise des structures de la langue ou 31 items de compréhension écrite. Les items ont été ordonnés de façon croissante selon leur difficulté [-3,94; 3,07], mesurée par l'application du modèle de Rasch à un paramètre pour items dichotomiques. Sans connaître leur valeur de difficulté, les experts ont estimé les limites des niveaux 3/4 à 10/11 selon les NCLC sur les items du TCF, avec la méthode du marque-page et en fixant le seuil de maîtrise des candidats à 66,67 %. Deux tours de jugement, pour chaque dimension, ont été nécessaires. Compte tenu que cette référentialisation s'est effectuée entièrement à distance, cette communication discute, de façon critique, les aspects méthodologiques et les principaux résultats.

BLOC K | Vendredi 12h15-12h45

K1 **Regard sur les pratiques d'évaluation des enseignantes du préscolaire à partir d'un questionnaire autorapporté**

Monica **Boudreau** et Marie-Hélène **Hébert** (UQAR)

Avec l'arrivée de la notion de compétence au préscolaire, les pratiques d'évaluation des enseignantes ont dues être revisitées. À quoi ressemblent-elles désormais ? Sont-elles en adéquation avec les prescriptions du ministère de l'Éducation en matière d'évaluation (MELS, 2011, 2012; MEQ, 2001, 2002, 2003) ? C'est dans ce contexte qu'ont été remises en question, au moyen d'un questionnaire auto-rapporté, les pratiques évaluatives d'enseignantes du préscolaire de quatre commissions scolaires du Québec, ainsi que leurs besoins de formations et de documentation dans le domaine de l'évaluation. D'abord validé par un comité d'experts de forme et de contenu, le questionnaire a été rempli à l'hiver 2013 par 106 enseignantes du préscolaire. Parmi les résultats préliminaires, il appert que les enseignantes ont surtout recours aux tâches/activités, à l'observation spontanée, aux grilles d'observation et au portfolio pour recueillir des données afin de porter un jugement sur le niveau de développement des compétences des enfants. Elles réclament, en outre, une plus grande disponibilité d'outils d'évaluation en tout genre pour optimiser leur compétence à évaluer et des formations liées aux thématiques suivantes : 1) les six compétences du programme d'éducation préscolaire (par exemple : Que signifient-elles ? À quoi les reconnaît-on chez les élèves ?) et 2) la démarche d'évaluation et ses étapes (par exemple : Comment consigner et « compiler » des traces qui serviront à la production du bulletin ? Comment uniformiser la notation d'une enseignante à l'autre ?).

K2 **Élaboration et de validation d'un modèle d'épreuve certificative en lecture et écriture axée sur l'approche par compétences en fin primaire au Sénégal : Problématique**

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

Le Sénégal, à l'instar de plusieurs pays d'Europe, d'Amérique ou d'Afrique, a emprunté le virage du renouveau de ses programmes scolaires pour les arrimer à la perspective de l'approche par compétences. Certains y voient une volonté de répondre aux besoins et préoccupations de la société sénégalaise, d'autres un suivisme à la mode imposé par les partenaires financiers à l'heure de la mondialisation. Au-delà du regard normatif et prescriptif des modèles, leur contextualisation en regard des enjeux sociaux s'impose : Faudrait-il continuer à reproduire des modèles extérieurs et à percevoir l'évaluation comme outil de sanction ou plutôt tenir compte, dans les épreuves, des caractéristiques de certaines disciplines, principalement le français, et du contexte local ?

Dans cette communication, l'évolution du système éducatif sénégalais, et spécifiquement de l'évaluation certificative en fin de primaire au Sénégal, sera décrite sous l'angle de la justice, de l'égalité et de l'équité. Également, des modèles de situations expérimentés dans différents milieux, notamment ceux réalisés en Europe et au Canada, seront analysés et critiqués au regard de la grille de Durand et Ducharme (2012) et du contexte sénégalais où le français, langue d'enseignement, est une langue seconde. Des recommandations seront faites et constitueront une sorte de cahier de charges qui va inspirer le modèle que nous allons élaborer. Le modèle de recherche développement de Harvey et Loiselle (2009) servira à la validation du prototype construit. Les différentes validations se feront à partir de mises à l'essai par divers acteurs sénégalais (enseignants et enseignantes, inspecteurs fonctionnaires des directions) impliqués dans la production d'épreuves certificatives.

K3

Analyse factorielle exploratoire d'un questionnaire sur les pratiques d'évaluation en salle de classe

Nadine **Talbot**, Gilles **Raïche** et Nathalie **Michaud** (UQAM)

Depuis les dernières réformes scolaires québécoises, certains auteurs ont écrit sur des caractéristiques de l'évaluation des apprentissages en salle de classe. Parmi celles-ci se retrouvent, notamment, la complexité, l'authenticité, le besoin de plusieurs critères ainsi que l'intégration de l'évaluation à l'enseignement. Les écrits publiés jusqu'à maintenant se concentrent principalement sur l'aspect théorique des caractéristiques et comment chacune d'elles pourraient se refléter dans des modalités ou des tâches d'évaluation des apprentissages en salle de classe. Conséquemment, peu d'outils ont été élaborés pouvant servir dans le cadre de recherche sur les pratiques d'évaluation des apprentissages en salle de classe. En accord avec son cadre de référence développé en 2006, Raïche et collab. (non publié) ont développé un questionnaire visant à établir si les quatre dimensions théoriques soit: planification, intégration, équité et authenticité, sont présentes dans les pratiques des enseignants. Au cours de cette contribution, une analyse factorielle exploratoire de ce questionnaire sera présentée.

K4

Méthodes quantitatives d'estimation des différences entre les hommes et les femmes en sciences sociales

Caroline **Simard** et Claudia **Savard** (Université Laval)

Dans l'ensemble des recherches scientifiques portant sur des concepts psychologiques, il est reconnu sans équivoque que le sexe du participant constitue une variable modératrice incontournable. En effet, sur une multitude d'aspects, les hommes et les femmes s'avèrent différents. Lorsque les analyses portent sur des données quantitatives, plusieurs techniques éprouvées sont disponibles afin de déterminer l'ampleur et la nature de ces différences. Les approches se différencient principalement l'une de l'autre de par leur niveau de complexité, leur précision et la quantité d'information qu'elles fournissent. Les principales méthodes sont : le test de Student (Abramowitz & Stegun, 1965), communément appelé le test t, qui compare les moyennes entre deux groupes; l'analyse de covariance (ANCOVA; voir Tabachnick & Fidell, 2012), qui estime l'impact du sexe en tant que covariable; l'analyse d'invariance, qui compare des modèles d'équations structurelles pour les deux groupes (Byrne, 2006); l'analyse du fonctionnement différentiel de l'item (DIF), qui sert à comparer la façon dont les hommes et les femmes répondent aux mêmes questions d'une mesure (Camilli & Shepard, 1994); et enfin, le modèle d'interdépendance acteur-partenaire (APIM; Kashy & Kenny, 1999), qui permet d'étudier l'interdépendance entre les hommes et les femmes d'une même dyade ou les deux partenaires d'un couple. En prenant exemple sur des données d'études, chacune de ces techniques sera présentée.

Formulaire d'inscription

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

Les frais d'inscription incluent le cocktail du jeudi soir et le dîner du vendredi midi.

Le dîner du jeudi n'est pas inclus.

J'assisterai au cocktail du jeudi : OUI NON

J'assisterai à l'Assemblée générale de l'ADMEE du jeudi OUI NON

J'assisterai au dîner du vendredi : OUI NON

Date limite de réception des formulaires d'inscription : **5 novembre 2013**

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

Inscription au congrès 300 \$

Inscription au congrès (ÉTUDIANT à temps complet) 150 \$ (Veuillez préciser votre université d'appartenance à INSTITUTION)

Transmettez-nous maintenant votre formulaire d'inscription électronique par courriel et ensuite par la poste avec votre chèque.

Cochez si : Adresse personnelle Adresse au bureau

Cochez si vous êtes présentatrice ou présentateur

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

Pour chaque bloc de communications, cocher la case de votre 1^{er} choix et celle de votre 2^e choix.

Jeudi 14 novembre 2013

Bloc A (13h15-13h45)		Bloc B (13h45-14h15)		Bloc C (14h30-15h)		Bloc D (15h-15h30)	
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"/>
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"/>
A3 <input type="checkbox"/>	A3 <input type="checkbox"/>	B3 <input type="checkbox"/>	B3 <input type="checkbox"/>	C3 <input type="checkbox"/>	C3 <input type="checkbox"/>	D3 <input type="checkbox"/>	D3 <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"/>
Bloc E (15h45-16h15)		Bloc F (16h15-16h45)		Bloc G (17h-17h30)		Bloc H (17h30-18h)	
1 ^{er} choix	2 ^e choix	1 ^{er} choix	2 ^e choix	1 ^{er} choix	2 ^e choix	1 ^{er} choix	2 ^e choix
E1 <input type="checkbox"/>	E1 <input type="checkbox"/>	F1 <input type="checkbox"/>	F1 <input type="checkbox"/>	G1 <input type="checkbox"/>	G1 <input type="checkbox"/>	H1 <input type="checkbox"/>	H1 <input type="checkbox"/>
E2 <input checked="" type="checkbox"/>	E2 <input checked="" type="checkbox"/>	F2 <input type="checkbox"/>	F2 <input type="checkbox"/>	G2 <input type="checkbox"/>	G2 <input type="checkbox"/>	H2 <input type="checkbox"/>	H2 <input type="checkbox"/>
E3 <input type="checkbox"/>	E3 <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"/>
E4 <input type="checkbox"/>	E4 <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"/>

Vendredi 15 novembre 2013

Bloc I (11h-11h30)		Bloc J (11h30-12)		Bloc K (12h15-12h45)	
1 ^{er} choix	2 ^e choix	1 ^{er} choix	2 ^e choix	1 ^{er} choix	2 ^e choix
I1 <input type="checkbox"/>	I1 <input type="checkbox"/>	J1 <input type="checkbox"/>	J1 <input type="checkbox"/>	K1 <input type="checkbox"/>	K1 <input type="checkbox"/>
I2 <input type="checkbox"/>	I2 <input type="checkbox"/>	J2 <input type="checkbox"/>	J2 <input type="checkbox"/>	K2 <input type="checkbox"/>	K2 <input type="checkbox"/>
I3 <input type="checkbox"/>	I3 <input type="checkbox"/>	J3 <input type="checkbox"/>	J3 <input type="checkbox"/>	K3 <input type="checkbox"/>	K3 <input type="checkbox"/>
I4 <input type="checkbox"/>	I4 <input type="checkbox"/>	J4 <input type="checkbox"/>	J4 <input type="checkbox"/>	K4 <input type="checkbox"/>	K4 <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 /

La section suivante doit être complétée lorsque les frais d'inscription sont défrayés par votre institution :

Nom de l'organisme :			
Personne-ressource :			
Courriel :		Téléphone :	