



Association pour le développement des méthodologies d'évaluation en éducation

---

## 31<sup>e</sup> session d'étude de l'ADMEE-Canada

# Les politiques et leur impact sur les pratiques évaluatives

Les 26 et 27 novembre 2009



le château bonne entente  
tradition & tendance

## Mot de bienvenue

C'est avec beaucoup d'enthousiasme que nous vous invitons à participer à la 31<sup>e</sup> session d'études de l'*Association pour le développement des méthodologies d'évaluation en éducation (ADMEE)*, qu'il nous importe d'ancrer dans une dynamique participative et conviviale. Parler et échanger nos visions des choses avec des collègues est un moyen d'examiner ensemble la signification des réalités relatives à l'acte d'évaluer. L'édition de cette année se tiendra à Québec les 26 et 27 novembre et j'invite tous les professionnels de l'éducation et de l'évaluation à se joindre à nous. Deux ateliers et plusieurs conférences vous seront offerts. Ce congrès visera à favoriser le partage d'expérience entre chercheurs, praticiens, professionnels et étudiants. C'est une occasion unique de mettre à jour vos connaissances professionnelles, de comparer vos expériences avec celles de vos collègues et de développer de nouveaux projets.

Cette année, l'ADMEE propose aux participantes et aux participants de porter un regard sur l'impact des politiques sur les pratiques évaluatives. La conférence d'ouverture sera prononcée par monsieur Pierre Brochu du Conseil des ministres de l'éducation du Canada et montrera les possibilités en matière d'analyse qu'offrent les évaluations à grande échelle administrées au Canada pour éclairer les politiques servant aux évaluations. M. Brochu présentera également les limites de ces évaluations dans leur « pouvoir explicatif » et les pratiques exemplaires dans le domaine.

En terminant, je vous souhaite un congrès stimulant et enrichissant. J'espère que vos échanges seront des plus fructueux et qu'ils ouvriront des perspectives fécondes pour l'avenir. Au plaisir de vous rencontrer à Québec les 26 et 27 novembre prochains et de partager avec vous ces bons moments.

Bonne session d'étude!

Pierre Valois  
Président de l'ADMEE Canada

## Jeudi 26 novembre 2009

9 h 00	–	10 h 00	Accueil et inscription
10 h 00	–	11 h 30	Ateliers 1 et 2
11 h 30	–	13 h 00	Dîner libre
13 h 00	–	13 h 30	Communications Bloc A
13 h 45	–	14 h 15	Communications Bloc B
14 h 30	–	15 h 00	Communications Bloc C
15 h 00	–	15 h 30	Pause
15 h 30	–	16 h 00	Communications Bloc D
16 h 15	–	16 h 45	Communications Bloc E
16 h 45	–	18 h 00	Cocktail
18 h 00	–	20 h 00	Assemblée générale de l'ADMEE

### ATELIERS - (JEUDI de 10 h à 11 h 30)

- Atelier 1 **Quels modèles psychométriques pour les politiques d'évaluation locale?**  
*Personnes-ressources :* Réginald Burton et Christophe Dierendonck (Université du Luxembourg)
- Atelier 2 **L'utilisation de la technologie dans un contexte d'évaluation certificative : expérimentation de la passation en ligne d'une épreuve d'écriture en français**  
*Personnes-ressources :* Jean-Guy Blais (Université de Montréal), Martin Paré (Mezur.com), Linda Drouin et Louise Pintal (MELS)

### BLOC A - (JEUDI de 13 h à 13h30)

- A1 **La politique de reconnaissance des acquis et des compétences des camionneurs québécois et ses impacts sur le secteur industriel, l'entreprise, l'individu et les pratiques d'évaluation des établissements d'enseignement : le cas des CFP en transport**  
*Conférencier et conférencières :* Jean Blanchet (Centre de formation professionnelle en transport de Charlesbourg), Lucie Héon et Liette Goyer (Université Laval)

- A2 **Les principaux aspects de l'intégration des pratiques d'évaluation des apprentissages aux pratiques pédagogiques**  
*Conférencières et conférencier :* Diane Leduc, Hélène Meunier, Nadine Talbot et Gilles Raïche (UQAM)
- A3 **Présentation d'un outil logiciel pour soutenir la détection des tentatives de fraudes lors de l'évaluation des apprentissages**  
*Conférenciers :* Sébastien Béland, Gilles Raïche (UQAM), Davis Magis (KU Leuven, Belgique) et Martin Lesage (UQAM)

### BLOC B - (JEUDI de 13 h 45 à 14 h 15)

- B1 **L'entente France-Québec sur la reconnaissance mutuelle des qualifications : un regard sur les instruments de mesure et d'évaluation des acquis de l'expérience en Europe**  
*Conférencier :* Michel Lejeune (Institut de recherche scientifique, Université du Québec)
- B2 **La qualité des évaluations dans un contexte de prépondérance du jugement**  
*Conférencier :* Pascal Ndinga (UQAM)
- B3 **L'apport de la modélisation de type Rasch pour mieux cerner les différents niveaux d'habiletés scientifiques chez les jeunes de 13 à 16 ans**  
*Conférenciers :* Éric Dionne (Université d'Ottawa) et Jesus Vazquez-Abad (Université de Montréal)

### BLOC C - (JEUDI de 14 h 30 à 15 h)

- ~~C1 **Politiques et pratiques de FTLV : quelle(s) évaluation(s) pour un système durable?** ANNULÉE~~  
~~*Conférencier et conférencière :* Gilles Pinte et Roseline Le Squere (Université de Bretagne Sud)~~
- C2 **Évaluation d'écoles : quel apport pour l'affirmation de l'autonomie des écoles et de leurs acteurs**  
*Conférencières :* Maria Manuela Martins Alves Terrasêca, Anabela Barros Pinto Sousa et Carina Luisa Moreira Coelho (Université de Porto)

- C3 DichoDif : un package du logiciel R pour la détection du fonctionnement différentiel d'items dichotomiques en psychométrie**  
*Conférenciers :* David Magis (KU Leuven), Sébastien Béland, (UQAM), Francis Tuerlinckx, Paul De Boeck (KU Leuven) et Gilles Raïche (UQAM)

#### **BLOC D - (JEUDI de 15 h 30 à 16 h)**

- D1 Utilisation de questionnaires Web ou téléphonique, quels sont les avantages et les inconvénients?**  
*Conférencières et conférencier :* Jacinthe Cloutier, Éric Frenette et Marie J. Lachance (Université Laval)
- D2 Le potentiel diagnostique des épreuves standardisées au Luxembourg**  
*Conférenciers et conférencière :* Christophe Dierendonck, Réginald Burton et Annick Fagnant (Université du Luxembourg)
- D3 Utilisation du Mini-Entrevues Multiples en contexte francophone : étude de généralisabilité**  
*Conférencière et conférenciers :* Christina St-Onge, Daniel J. Côté (Université de Sherbrooke) et Carlos Brailovsky (Université Laval)

#### **BLOC E - (JEUDI de 16 h 15 à 16 h 45)**

- E1 L'évaluation fonctionnelle des troubles graves de comportement : proposition d'une procédure basée sur les données probantes**  
*Conférencière et conférenciers :* MéliSSa Gaucher (UQAM), Michel Rolland (Hôpital Rivière-des-Prairies) et Jacques Forget (UQAM)
- E2 Un modèle d'évaluation certificative cohérent avec la Politique québécoise d'évaluation des apprentissages : comment créer des liens entre les intentions politiques et la salle de classe**  
*Conférencier :* Éric Dionne (Université d'Ottawa)
- E3 L'importance de la spécificité des rôles professionnels lors de l'évaluation de la perception d'efficacité personnelle des directions d'école**  
*Conférenciers et conférencière :* Claude Fernet (UQTR), Stéphanie Austin (Université Laval) et Marc Dussault (UQTR)

### **Vendredi 27 novembre 2009**

<b>8 h 30 – 9 h 00</b>	<b>Inscription</b>
<b>9 h 00 – 10 h 15</b>	<b>Conférence d'ouverture</b>
<b>10 h 15 – 10 h 30</b>	<b>Pause</b>
<b>10 h 30 – 11 h 00</b>	<b>Communications Bloc F</b>
<b>11 h 15 – 11 h 45</b>	<b>Communications Bloc G</b>
<b>11 h 45 – 13 h 30</b>	<b>Dîner de l'ADMEE</b>
<b>13 h 30 – 14 h 00</b>	<b>Communications Bloc H</b>
<b>14 h 15 – 14 h 45</b>	<b>Communications Bloc I</b>
<b>15 h 00 – 15 h 30</b>	<b>Communications Bloc J</b>

#### **CONFÉRENCE D'OUVERTURE (VENDREDI de 9 h à 10 h 15)**

- Devrait-on se fonder sur les évaluations à grande échelle pour orienter les politiques en éducation?**  
*Conférencier :* Pierre Brochu, Conseil des ministres de l'éducation du Canada

#### **BLOC F - (VENDREDI de 10 h 30 à 11 h)**

- F1 Évolution de l'évaluation des apprentissages au Québec du début du XX<sup>e</sup> siècle à nos jours : politiques et pratiques évaluatives**  
*Conférenciers et conférencière :* Adrien Cantat, Pierre Valois (Université Laval) et Marie-Hélène Hébert (UQAR)
- F2 Expertest, une percée dans le milieu des entreprises et des corporations**  
*Personnes-ressources :* Yves Groulx et Jacques Bellemare (Société GRICS)
- F3 L'évaluation authentique en contexte d'éducation physique au primaire**  
*Conférencière et conférenciers :* Johanne Grenier, Gilles Raïche (UQAM), Pierre Boudreau (Université d'Ottawa) et Jean-François Richard (Université de Moncton)

#### **BLOC G - (VENDREDI de 11 h 15 à 11 h 45)**

- G1 **L'intégration des pratiques d'évaluation des apprentissages aux pratiques pédagogiques dans le contexte des approches par compétences**  
*Conférencières et conférencier* : Diane Leduc, Hélène Meunier, Nadine Talbot et Gilles Raïche (UQAM)
- G2 **Les politiques d'évaluation de l'écriture et leur apport quant aux pratiques éducatives**  
*Conférencier et conférencière* : Serge Sévigny (Université Laval) et Isabelle Beaudoin (UQAR)
- G3 **Identification des différences sur le plan des composantes de la violence selon la position sur le jeu au soccer, le niveau de jeu et le rôle du joueur**  
*Conférenciers* : Julien D'Amours-Raymond, Éric Frenette (Université Laval) et Martin Gendron (UQAR)

#### **BLOC H - (VENDREDI de 13 h 30 à 14 h)**

- ~~H1 **La différence entre la correction de textes manuscrits et de textes dactylographiés, puis imprimés à l'ordinateur**~~  
~~*Conférencière et conférenciers* : Caroline Godin, Jean-Guy Blais et Michel Laurier (Université de Montréal)~~ **ANNULÉE**
- H2 **L'évaluation des politiques institutionnelles d'évaluation des apprentissages (PIEA) des collèges du Québec**  
*Conférencière* : Katie Bérubé (Commission d'évaluation de l'enseignement collégial)
- H3 **Le jugement crédible en évaluation de programme**  
*Conférencière et conférencier* : Marthe Hurteau (UQAM) et Pierre Valois (Université Laval)

#### **BLOC I - (VENDREDI de 14 h 15 à 14 h 45)**

- I1 **Autoévaluation d'une épreuve adaptative à deux niveaux pour l'évaluation du rendement scolaire des élèves dans le cadre d'enquêtes à grande échelle**  
*Conférencières* : Julie Grondin et Nancy Boiteau (UQAR)
- I2 **L'utilisation des politiques institutionnelles d'évaluation des apprentissages pour assurer la validité des inférences d'évaluation des apprentissages**  
*Conférencière* : Karine Paquette-Côté (UQAM)
- I3 **Évaluation des cours de la M.Éd entièrement en ligne en administration scolaire et en enseignement, apprentissage et évaluation à l'Université d'Ottawa**  
*Conférencières et conférencier* : Hélène Meunier (UQAM), Claire Isabelle, Emmanuel Duplâa (Université d'Ottawa) et Hélène Fournier (CNRC)

#### **BLOC J- (VENDREDI de 15 h à 15 h 30)**

- J1 **L'apport des correcticiels dans la correction de textes d'élèves du secondaire**  
*Conférencières et conférencier* : Marie Hélène Mireault, Jean-Guy Blais et Marie-Claude Boivin (Université de Montréal)
- J2 **Évaluation des attitudes en stage au collégial**  
*Conférencière et conférencier* : Raymonde Gosselin (Collège Édouard-Montpetit) et Gilles Raïche (UQAM)
- J3 **L'apprentissage universitaire et la réussite dans le travail : une analyse sur le contexte, les résultats, les compétences acquises et le succès dans la carrière**  
*Conférencières et conférencier* : Carla Barroso da Costa, Nathalie Loye et Jake Murdoch (Université de Montréal)

## Jeudi 26 novembre 2009

9 h 00	–	10 h 00	Accueil et inscription
10 h 00	–	11 h 30	Ateliers 1 et 2
11 h 30	–	13 h 00	Dîner libre
13 h 00	–	13 h 30	Communications Bloc A
13 h 45	–	14 h 15	Communications Bloc B
14 h 30	–	15 h 00	Communications Bloc C
15 h 00	–	15 h 30	Pause
15 h 30	–	16 h 00	Communications Bloc D
16 h 15	–	16 h 45	Communications Bloc E
16 h 45	–	18 h 00	Cocktail
18 h 00	–	20 h 00	Assemblée générale de l'ADMEE

### ATELIERS - (JEUDI de 10 h à 11 h 30)

#### Atelier 1 **Quels modèles psychométriques pour les politiques d'évaluation locale?**

*Personnes-ressources :* Réginald Burton et Christophe Dierendonck (Université du Luxembourg), Lucie M-L

De plus en plus de dispositifs d'évaluation locale (évaluation formative, *monitoring*, épreuves d'orientation scolaire) ont recours aux Modèles de Réponse à l'Item pour estimer les compétences scolaires des élèves au départ de tests papier-crayon ou assistés par ordinateur. Ces modèles de mesure – dont les propriétés laissent entrevoir des progrès et des avantages considérables – se sont d'abord imposés comme standards de référence dans les enquêtes internationales pour être ensuite transposés dans des contextes locaux.

Cependant, la technicité des MRI, leur nature stochastique et leurs composantes statistiques relativement complexes n'ont pas toujours permis une approche cartésienne et prudente de leur exploitation. On a cru pendant longtemps que la transposition de ces modèles des enquêtes internationales aux évaluations locales était naturelle et ne souffrait d'aucun obstacle fondamental. Mais les objectifs, les caractéristiques et les exigences des évaluations locales sont parfois bien éloignés des préoccupations internationales, ce qui peut conduire à des erreurs de mesure.

Nous nous proposons donc, au départ de résultats empiriques, de déterminer dans quelles conditions l'application des MRI aux politiques d'évaluations locales est pertinente? Dans cette perspective, plusieurs critères d'analyse seront envisagés : la nature théorique des modèles (les MRI sont-ils adaptés pour atteindre les objectifs des évaluations locales?), les caractéristiques psychométriques des évaluations (les conditions d'application des MRI sont-elles toujours remplies dans le cadre des évaluations locales?) et les méthodes validation des modèles (les méthodes pour évaluer l'adéquation des modèles sont-elles adaptées aux évaluations locales?). Pour finir, nous nous interrogerons sur l'opportunité de développer de nouveaux modèles pour l'évaluation des compétences.

#### Atelier 2 **L'utilisation de la technologie dans un contexte d'évaluation certificative : expérimentation de la passation en ligne d'une épreuve d'écriture en français**

*Personnes-ressources :* Jean-Guy Blais (Université de Montréal), Martin Paré (Mezur.com), Linda Drouin et Louise Pintal (MELS)

Avec l'utilisation croissante des technologies de l'information et de la communication dans les écoles (ordinateurs et Internet) le ministère de l'Éducation, des Loisirs et des Sports du Québec (MELS) est souvent interpellé par les établissements scolaires sur des questions liées à l'utilisation de ces technologies dans le contexte de la passation des épreuves ministérielles. Des demandes insistantes sont formulées par certains établissements qui ont mis de l'avant dans leur projet éducatif le recours aux technologies. Ces milieux souhaiteraient notamment que leurs élèves puissent passer les épreuves ministérielles au moyen du traitement de texte en situation d'évaluation de la compétence à écrire. Par souci d'équité et de justice à l'endroit des élèves, le MELS s'interroge sur les conditions et conséquences du recours aux technologies pour la passation d'épreuves certificatives en ligne. En collaboration avec des chercheurs du GRIÉMÉtic de l'Université de Montréal, une expérience a donc été réalisée en ce sens auprès de 65 élèves de quatrième secondaire d'une école de la région de Québec en mai 2009. La présentation fera état des caractéristiques de la plateforme de passation en ligne développée à cet effet, des observations réalisées durant la passation et des commentaires des enseignants et des élèves au sujet de leur expérience.



L'expérience réalisée permet de fournir des éléments de réponse aux questions suivantes :

- Quels sont les avantages et les inconvénients de l'utilisation des technologies en contexte d'évaluation certificative de la compétence à écrire?
- Dans une perspective de moyen et de long terme, quelle démarche le MELS devrait-il privilégier pour être en mesure d'assurer le développement d'épreuves certificatives qui tiennent compte des avancées relatives à l'utilisation des technologies de l'information et de la communication dans les écoles?

## **BLOC A - (JEUDI de 13 h à 13h30)**

### **A1 La politique de reconnaissance des acquis et des compétences des camionneurs québécois et ses impacts sur le secteur industriel, l'entreprise, l'individu et les pratiques d'évaluation des établissements d'enseignement : le cas des CFP en transport**

*Conférencier et conférencières* : Jean Blanchet (Centre de formation professionnelle en transport de Charlesbourg), Lucie Héon et Liette Goyer (Université Laval)

Au Québec, dans la foulée de la *Politique gouvernementale d'éducation des adultes et de formation continue* (2002), il fut proposé un Cadre général et technique (2005) qui est venu préciser les valeurs et la démarche afférentes à la reconnaissance des acquis et des compétences (RAC) en formation professionnelle et technique. Dans le but annoncé d'élargir les services de RAC, le ministère de l'Éducation, du Loisir et du Sport (MELS) a relayé aux établissements une large responsabilité quant à l'organisation des services et au développement des politiques et pratiques d'évaluation. Faisant appel à l'adaptabilité des organisations concernées, le MELS les prévenait que « les activités d'évaluation à mettre en place [...] doivent être rigoureuses, fiables et assorties de modalités d'évaluation adaptées à la nature extrascolaire et au caractère généralement expérientiel des apprentissages réalisés par la personne » (Cadre général, p. 5).

À titre de conseiller pédagogique responsable des services de RAC dans l'un des deux centres nationaux de formation professionnelle dédiés au transport, Jean Blanchet côtoie les intervenants et entreprises du domaine du transport routier qui, depuis deux ans, exercent une forte demande en

matière de qualification professionnelle des camionneurs. Voici donc, du point de vue du praticien, les principaux résultats de l'analyse des impacts de cette nouvelle approche d'évaluation.

L'analyse des retombées des services de RAC dans le domaine du transport sur les plans du secteur industriel (macro), des entreprises (méso), des individus (micro) ainsi que des établissements d'enseignement et de leurs politiques d'évaluation (endo) est réalisée à l'aide de la « matrice d'indicateurs de l'impact des formations qualifiantes et transférables » proposée par Paul Bélanger (2004).

### **A2 Les principaux aspects de l'intégration des pratiques d'évaluation des apprentissages aux pratiques pédagogiques**

*Conférencières et conférencier* : Diane Leduc, Hélène Meunier, Nadine Talbot et Gilles Raïche (UQAM)

L'intégration des pratiques d'évaluation des apprentissages aux pratiques pédagogiques est une préoccupation importante pour tout professeur qui a à cœur la réussite de ses étudiants. Celui-ci enseigne dans le but de faire réussir les tâches d'évaluation. Il cherche à leur donner un sens, autant pour lui, qui les crée, que pour ceux qui sont évalués. De plus, lorsqu'il est question d'intégration, il est aussi question de modalités d'évaluation en parfaite harmonie avec les objectifs visés et en constante évolution. Ainsi, dans une approche par compétence, chacune de ces modalités devrait permettre à l'étudiant de se préparer à l'évaluation suivante. Enfin, l'intégration des pratiques d'évaluation des apprentissages aux pratiques pédagogiques touche au programme d'étude et est l'un des porte-étendard de l'approche programme dans laquelle les apprentissages sont évalués graduellement, de façon cohérente, tout au long du parcours de l'étudiant.

Ces trois aspects de l'intégration des pratiques d'évaluation des apprentissages aux pratiques pédagogiques seront abordés en lien avec la dernière politique d'évaluation des apprentissages du ministère de l'Éducation du Québec (Gouvernement du Québec, 2003). Nous mettrons en évidence les extraits de cette politique qui touchent, directement ou indirectement, l'intégration des pratiques et nous tenterons de cerner, s'il y a lieu, les problématiques qu'elle sous-tend.

### A3 **Présentation d'un outil logiciel pour soutenir la détection des tentatives de fraudes lors de l'évaluation des apprentissages**

*Conférenciers* : Sébastien Béland, Gilles Raïche (UQAM), Davis Magis (KU Leuven, Belgique) et Martin Lesage (UQAM)

Lors de l'administration d'instruments visant à mesurer le niveau d'habileté d'un étudiant, il est fréquent d'assister à des tentatives de fraude. Évidemment, ce genre de problème peut être minimisé par l'intermédiaire de stratégies de prévention que les administrateurs et les enseignants connaissent généralement bien (nous pensons, par exemple, à un local bien éclairé ou à recourir à plusieurs surveillants). Or, une approche plus sophistiquée consiste à utiliser des indices statistiques permettant de détecter les étudiants qui présentent des patrons de réponses inappropriés.

Dans le cadre de ce colloque, nous présenterons les constituantes d'une toute nouvelle librairie disponible via le logiciel R : irtFit. Cet outil logiciel inclut une vingtaine d'indices de détection paramétriques et non paramétriques pouvant être appliqués à des matrices de réponses dichotomiques. À la marge, notre présentation sera articulée autour de trois grands pôles : après un bref survol du fonctionnement de notre outil logiciel, nous offrirons un exemple d'application permettant de mettre en exergue son champ d'application. Enfin, nous discuterons des ouvertures relatives au développement d'une deuxième version de la librairie irtFit.

## **BLOC B - (JEUDI de 13 h 45 à 14 h 15)**

### B1 **L'entente France-Québec sur la reconnaissance mutuelle des qualifications : un regard sur les instruments de mesure et d'évaluation des acquis de l'expérience en Europe**

*Conférencier* : Michel Lejeune (Institut de recherche scientifique, Université du Québec)

La prise en compte de ce qui se fait en Europe et des progrès dans le domaine de l'évaluation pourrait permettre d'accélérer le développement des mécanismes de reconnaissance des expériences et des compétences, conformément à l'entente France-Québec sur la reconnaissance mutuelle des qualifications. Les instruments de reconnaissance des acquis de l'expérience constituent d'ailleurs un des principaux enjeux du Cadre européen de certifications professionnelles auquel se réfèrent les politiques et les organismes français spécialisés dans la qualification professionnelle.

Les expériences européennes font l'objet d'une veille technologique qui s'alimente d'information provenant de pays qui ont adopté la reconnaissance des apprentissages non formels et informels. L'Union européenne, au même titre que la France et le Québec, admet volontiers les types d'instruments suivants : le test traditionnel, les techniques déclaratives (incluant le portfolio), l'observation, la simulation, les évidences physiques et intellectuelles ainsi que l'autoévaluation. Cependant, la correspondance entre les instruments d'évaluation issus de l'éducation et les instruments propres au milieu de travail alimente quelques débats au sein de la communauté : d'aucuns remettent en cause la pertinence de certains instruments de mesure et d'évaluation en regard de la nature informelle de l'expérience et de son contexte d'action. D'ailleurs, l'influence de l'Angleterre, avec son cadre national de qualification (NVQ), semble déterminante. Ce modèle permettrait en effet de cerner différents aspects de l'expérience, à partir de sources multiples et diversifiées (différentes mesures et, ou, plusieurs experts ou observateurs). Cette approche fait lentement sa place dans les branches professionnelles en France, sous l'appellation d'« analyse plurimodale ».

Cette communication propose un tour d'horizon des pratiques innovantes, ou émergentes, propres à l'Europe, en matière d'évaluation des acquis de l'expérience (possibilités et limites), en regard de l'entente France-Québec sur la reconnaissance mutuelle des compétences des travailleurs qualifiés.

### B2 **La qualité des évaluations dans un contexte de prépondérance du jugement**

*Conférencier* : Pascal Ndinga (UQAM)

Depuis de nombreuses années, des efforts appréciables ont été déployés pour assurer la qualité des évaluations réalisées dans les écoles du Québec et ailleurs dans le monde. Des sessions de formation comme celles organisées par l'ADMEE, des communications scientifiques, des colloques et des congrès internationaux témoignent de cette préoccupation constante quant à l'amélioration de la qualité des évaluations. L'implantation récente du renouveau pédagogique a engendré une politique fondée sur la prépondérance du jugement en évaluation. Cette innovation constitue en soi un défi colossal pour le monde de l'éducation en général, pour le domaine de la mesure et de l'évaluation en particulier. Le jugement, par sa nature propre, comporte une part non négligeable de subjectivité. Cette dernière



doit être dosée et quantifiée afin de cerner la part qui symbolise la qualité réelle du jugement. Les outils de jugement, largement déployés dans le milieu scolaire, ne sont guère rassurants quant à la précision de l'information qu'ils fournissent. De plus, très peu d'études sont réalisées sur l'évaluation en contexte de prédominance du jugement. Ainsi, comment assurer la qualité des évaluations? Cette question mérite une attention particulière compte tenu des multiples implications des décisions découlant des évaluations où le jugement prédomine.

**B3 L'apport de la modélisation de type Rasch pour mieux cerner les différents niveaux d'habiletés scientifiques chez les jeunes de 13 à 16 ans**

*Conférenciers* : Éric Dionne (Université d'Ottawa) et Jesus Vazquez-Abad (Université de Montréal)

L'approche par compétences pose de sérieux défis, notamment en ce qui concerne les pratiques évaluatives. Depuis la fin de l'année 2003, les intervenants québécois du primaire et du secondaire peuvent se référer à la *Politique d'évaluation des apprentissages* (Gouvernement du Québec, 2003) afin de baliser la démarche évaluative. Cependant, il reste beaucoup de ponts à créer entre les intentions ministérielles et les pratiques évaluatives sur le terrain. Notre présentation fait état d'une recherche visant à développer et valider un instrument permettant de mieux cibler le niveau des habiletés des processus scientifiques chez les jeunes au début du secondaire (13-16 ans). En science, les habiletés de processus représentent des ressources mobilisables de toute première importance dans l'exercice d'une compétence. En ce sens, la connaissance du niveau de maîtrise des habiletés de processus devient un enjeu important en particulier dans une perspective d'évaluation diagnostique. De plus, à notre connaissance il n'y a pas d'outils développés en langue française et adaptés au contexte québécois de l'enseignement des sciences.

Dans la présentation, nous exposerons les sept habiletés mesurées à savoir : (1) interpréter les données, (2) définir opérationnellement, (3) énoncer des hypothèses, (4) synthétiser l'information, (5) évaluer l'information trouvée, (6) planifier et organiser le travail et (7) s'entraider. Nous avons expérimenté l'outil auprès de 455 élèves du premier cycle du secondaire issus de deux commissions scolaires et de cinq écoles au début de l'année scolaire 2008-2009. Nous présenterons les résultats que nous avons obtenus

au moyen d'analyse de type Rasch. Les analyses démontrent que l'outil offre des perspectives intéressantes dans un contexte diagnostique par l'étude des leurres qui représentent tous un construit particulier en science.

**BLOC C - (JEUDI de 14 h 30 à 15 h)**

~~C1 Politiques et pratiques de FTLV : quelle(s) évaluation(s) pour un système durable?~~

~~Conférencier et conférencière : Gilles Pinte et Roseline Le Square  
(Université de Bretagne Sud) ANNULÉE~~

En France, les débats sur l'efficacité du système de formation deviennent récurrents depuis les années 1990. Plusieurs rapports ont mis en avant l'opacité financière du système et le manque de lisibilité du dispositif. Au cœur de cette problématique se trouve la formation continue (FC), l'une des missions principales des établissements d'enseignement supérieur (universités et écoles d'ingénieurs), avec plus de 300 000 stagiaires formés chaque année. Dans le même temps, la recomposition du champ de l'éducation des adultes a favorisé l'évolution des politiques et des pratiques de ce système de formation.

On a souvent demandé à la FC d'avoir une fonction d'infirmier adaptative, voire rééducative, mais aussi de donner une seconde chance aux décrocheurs de la formation initiale ou encore de favoriser la promotion sociale avec souvent des résultats très probants. Depuis les années 1980, on a demandé au système FC de reconverter et de requalifier des milliers de salariés devenus demandeurs d'emploi ou risquant de le devenir et d'organiser une liste d'attente pour une reprise du travail. Depuis les années 1990, on demande aux formateurs d'accompagner les transitions de carrière que l'on prévoit désormais discontinues et flexibles.

La FC, outil pour faire face aux défis de la compétence, assorti aux objectifs d'encouragement de l'inclusion sociale et de sécurisation des parcours, est incontesté en France, en Europe. Même si le débat sur l'efficacité d'un système est légitime, il pose également en contrepoint celui de l'évaluation.

Aussi, les auteurs se proposent de revoir les objectifs que l'on a assignés à la FC, d'en interroger les modes d'évaluation et de dresser quelques pistes d'élaboration de critères permettant d'aller vers une EFTLV (Éducation et formation tout au long de la vie) durable. Ils se saisiront également du

contexte d'une recherche européenne sur la mise en œuvre des politiques FTLV et leurs retombées.

## C2 **Évaluation d'écoles : quel apport pour l'affirmation de l'autonomie des écoles et de leurs acteurs**

*Conférencières* : Maria Manuela Martins Alves Terrasêca, Anabela Barros Pinto Sousa et Carina Luisa Moreira Coelho (Université de Porto)

Plus les États sont décentralisés, d'autant plus ils se préoccupent de la régulation des processus éducatifs.

Cette régulation n'est pas faite, dans ces cas-là, *a priori* (sous peine de se mettre en position de commander la politique au lieu de la décentraliser) mais elle est faite à travers la mise en œuvre de procédures de contrôle, désignées comme en étant des processus d'évaluation.

Cette circonstance crée, à notre avis, un ensemble de contraintes pour l'action des acteurs : d'une part, on leur donne la possibilité de s'autodéterminer et, d'autre part, on voit émerger un ensemble de conditionnements qui retombent sur les produits et qui empêchent que les pratiques évaluatives puissent, effectivement, s'ajuster au spécifique plutôt qu'à ce qui est singulier, pour pouvoir répondre aux singularités.

Dans un travail d'évaluation des écoles, au Portugal, il y a une pression, toute-puissante, de l'évaluation externe sur les établissements scolaires. Cette pression oriente – à travers des règles strictes, exigeantes et précises – les modes d'agir de l'évaluation interne. Cette orientation empêche les acteurs de réaliser une réflexion critique et approfondie sur eux-mêmes et sur leurs actions, qui leur permettrait de s'affirmer en tant qu'auteurs de leurs propres pratiques. De cette façon, les identités des agents éducatifs (y inclus les professeurs) se conforment à des « identités clonées » (plutôt qu'attribuées...), parce que construites suivant une imposition du dehors vers l'intérieur.

Si l'autoévaluation arrive à s'affirmer en tant qu'indépendante de cette pression de l'évaluation externe, elle peut réussir et échapper à cette détermination.

Ces réflexions ont été approfondies dans le contexte d'un projet de recherche où l'équipe de recherche universitaire travaille en profonde liaison avec des équipes d'autoévaluation d'écoles du nord du Portugal.

## C3 **DichoDif : un package du logiciel R pour la détection du fonctionnement différentiel d'items dichotomiques en psychométrie**

*Conférenciers* : David Magis (KU Leuven), Sébastien Béland, (UQAM), Francis Tuerlinckx, Paul De Boeck (KU Leuven) et Gilles Raïche (UQAM)

La détection du fonctionnement différentiel des items est un domaine de recherche majeur en psychométrie moderne. La détection de tels items et leur retrait du test proposé constituent des tâches nécessaires et cruciales pour la validité de l'interprétation des résultats d'apprenants soumis au test considéré. De nombreuses méthodes de détection du fonctionnement différentiel des items dichotomiques ont été proposées ces dernières décennies, et beaucoup d'entre elles disposent de programmes propres pour leur utilisation pratique. Néanmoins, très peu d'entre elles sont disponibles de façon totalement gratuite et en téléchargement libre.

L'objectif de cet exposé est de présenter un nouveau *package*, nommé *dichoDif*, rassemblant une dizaine de méthodes de détection du fonctionnement différentiel d'items à réponses dichotomiques. Ces méthodes reposent sur l'utilisation (ou non) de modèles de réponses à l'item, permettent de détecter le fonctionnement différentiel uniforme ou non uniforme, et peuvent être utilisées avec un groupe focal d'apprenants ou plus. Le *package* est utilisable gratuitement à l'aide du logiciel R. Nous mettrons l'accent sur son utilité pratique en analysant un exemple réel et en comparant succinctement les différentes méthodes disponibles.

## **BLOC D - (JEUDI de 15 h 30 à 16 h)**

### D1 **Utilisation de questionnaires Web ou téléphonique, quels sont les avantages et les inconvénients?**

*Conférencières et conférencier* : Jacinthe Cloutier, Éric Frenette et Marie J. Lachance (Université Laval)

L'utilisation d'Internet en tant que moyen de communication est maintenant chose commune. C'est pourquoi plusieurs firmes offrent la possibilité de sonder un panel de participants à l'aide du Web, plutôt que l'approche traditionnelle par téléphone. Plusieurs avantages sont offerts par cette méthode tels la flexibilité, les coûts moindres, la rapidité, la possibilité d'obliger les participants à répondre à une question, l'obtention des données

de différentes localisations et au moment qui convient au répondant (Ardalan, Ardalan & Coppage, 2009; Booth-Kewley, Edwards & Rosenfeld, 1992; Healey, Macpherson & Kuijten, 2005; Mitchell, Klein & Balloun, 1996). Du côté des inconvénients, puisqu'il s'agit d'un échantillonnage non aléatoire, il est impossible de calculer une marge d'erreur (Agence de la consommation en matière financière du Canada, 2008), souvent le taux de réponse est faible (Ardalan et al., 2009; Healey et al., 2005) et le développement du questionnaire est parfois coûteux. Bien que cette technologie soit répandue, il existe encore des personnes anxieuses face à celle-ci. De plus si des erreurs sont commises dans la programmation du logiciel, ceci peut résulter en une perte de données (Evan & Miller, 1996; Mitchell et al., 1996). Une autre appréhension majeure des chercheurs concerne la qualité des données. Plus précisément, obtient-on les mêmes résultats selon que les données sont collectées par Web ou par téléphone? À partir d'un modèle de socialisation appliqué à l'épargne (connaissances, perceptions et pratiques en matière d'épargne et d'investissement), ces deux approches seront comparées à partir d'analyses d'invariance par équations structurelles. Les données proviennent d'un sondage qui a été effectué au cours du printemps 2008 auprès de 400 sujets par téléphone et 966 sujets par Web. La présentation exposera le cadre théorique lié à la problématique du choix de la méthode de passation des questionnaires par Web ou par téléphone.

## D2 Le potentiel diagnostique des épreuves standardisées au Luxembourg

*Conférenciers et conférencière* : Christophe Dierendonck, Réginald Burton et Annick Fagnant (Université du Luxembourg)

Depuis trois ans, le ministère de l'Éducation luxembourgeois s'efforce de mettre en place un *monitoring* du système dans l'objectif d'améliorer l'efficacité et l'équité de l'enseignement. Dans ce cadre, des épreuves standardisées sont proposées dans les classes de troisième primaire et de troisième secondaire dans les domaines des mathématiques et des langues. À la suite de ces évaluations semi-externes, un double *feedback* est envoyé aux enseignants : l'un normatif (mais anonyme) permettant de situer le niveau des élèves par rapport à l'ensemble de la population scolaire et l'autre, critérié, informant de la maîtrise de certaines compétences et ressources. L'hypothèse forte qui préside à cette démarche de *feedback* est

que les enseignants, mieux informés des forces et des faiblesses de leurs élèves, adopteront diverses actions de régulation leur permettant de compenser d'éventuels déficits mis en lumière par les épreuves. L'objectif de notre communication est de discuter cette hypothèse forte à la lumière d'une étude exploratoire menée dans le domaine des mathématiques et visant à développer le potentiel diagnostique des épreuves standardisées en vue de donner aux enseignants un *feedback* qui dépasse le simple constat lié à la maîtrise ou non de certains contenus ou de certaines compétences pour s'orienter davantage vers une analyse des difficultés rencontrées.

## D3 Utilisation du Mini-Entrevues Multiples en contexte francophone : étude de généralisabilité

*Conférencière et conférenciers* : Christina St-Onge, Daniel J. Côté (Université de Sherbrooke) et Carlos Brailovsky (Université Laval)

Le succès académique antérieur s'avère le meilleur prédicteur du succès académique en médecine; toutefois cet indicateur n'informe en rien quant aux habiletés non cognitives des candidats. Le Mini Entrevues Multiples (MEM), une mesure d'habiletés non cognitives, démontre une bonne fidélité et validité prédictive. L'objectif de la présente étude était de mesurer la fidélité de cet outil lors de son administration dans le cadre du processus de sélection au doctorat en médecine de l'Université de Sherbrooke. Les résultats observés démontrent la fidélité de l'outil, et ce, dans un contexte d'administration différent des études précédentes.

## BLOC E - (JEUDI de 16 h 15 à 16 h 45)

### E1 L'évaluation fonctionnelle des troubles graves de comportement : proposition d'une procédure basée sur les données probantes

*Conférencière et conférenciers* : Mélissa Gaucher (UQAM), Michel Rolland (Hôpital Rivière-des-Prairies) et Jacques Forget (UQAM)

Il est souvent difficile d'intervenir auprès de jeunes présentant des troubles graves de comportement tels que l'automutilation et les conduites violentes envers autrui. Dans ce contexte, les études montrent l'importance de l'évaluation dans la mise en place de plans d'intervention efficaces. Cependant, dans beaucoup d'établissements publics, la procédure d'évaluation est minimaliste, le personnel préférant intervenir rapidement sur les comportements problématiques plutôt que d'établir de manière systématique une évaluation individualisée. Par conséquent, l'intervention

est souvent la même pour tous et se révèle peu efficace. Néanmoins, dans le cas des troubles graves du comportement, il existe un certain nombre de procédures d'évaluation s'appuyant sur des données probantes. La meilleure pratique consiste à déterminer la fonction des comportements inadéquats. Pour ce faire, des méthodes d'évaluation directes et indirectes sont proposées dans la documentation scientifique. Bien qu'elles soient souvent nécessaires à une compréhension juste de la fonction des comportements, les mesures directes sont rarement utilisées. La présentation portera sur l'évaluation fonctionnelle, l'analyse fonctionnelle, l'analyse fonctionnelle expérimentale et l'analyse fonctionnelle psychophysiologique. Compte tenu des contraintes du milieu scolaire, plusieurs de ces méthodes sont difficiles à appliquer. Une méthode adaptée à ce milieu sera donc présentée. Enfin, ces méthodes d'évaluation seront illustrées par des études de cas.

## E2 **Un modèle d'évaluation certificative cohérent avec la Politique québécoise d'évaluation des apprentissages : comment créer des liens entre les intentions politiques et la salle de classe**

*Conférencier* : Éric Dionne (Université d'Ottawa)

La *Politique québécoise d'évaluation des apprentissages* réaffirme de façon marquée l'importance du jugement professionnel dans le difficile exercice de l'évaluation des compétences par les enseignantes et les enseignants du primaire et du secondaire à la formation générale des jeunes. Ce jugement est d'autant plus difficile à exercer et à mettre en œuvre quand les enjeux sont importants, comme c'est le cas dans le contexte de l'évaluation certificative. Les enseignantes et les enseignants doivent alors appuyer leur jugement et répondre de leur jugement au besoin auprès de différents acteurs comme le ministère de l'Éducation, du Loisir et du Sport (MELS), les directions d'école, les parents ou les élèves. Dans son *Plan de mise en œuvre de la politique d'évaluation des apprentissages*, le MELS souligne qu'il encouragera les initiatives novatrices en évaluation des apprentissages mais force est de constater que ces initiatives sont encore relativement peu nombreuses. Dans le cadre de cette présentation, nous proposerons le fruit de notre réflexion sur un modèle d'évaluation certificative qui, d'une part, vise à rendre compte du niveau de développement de compétences et, d'autre part, est cohérent avec l'esprit et les principes énoncés dans la *Politique d'évaluation des apprentissages* du MELS. Notre exposé dressera

un portrait sommaire des grandes lignes de la Politique. Nous enchaînerons en présentant le modèle d'évaluation certificative que nous proposons et qui tient compte, par exemple, de l'approche par compétences, des cycles d'apprentissage et des principes généralement admis en évaluation des apprentissages. Finalement, nous conclurons en abordant les avantages et les limites d'un tel modèle ainsi que les perspectives de recherche que ce modèle suscite.

## E3 **L'importance de la spécificité des rôles professionnels lors de l'évaluation de la perception d'efficacité personnelle des directions d'école**

*Conférenciers et conférencière* : Claude Fernet (UQTR), Stéphanie Austin (Université Laval) et Marc Dussault (UQTR)

De récentes études révèlent que les directions d'école vivent de plus en plus d'insatisfactions et de difficultés professionnelles. Confrontés à une tâche de travail qui se complexifie et qui s'alourdit, plusieurs directeurs remettent régulièrement en question leur choix de carrière (FQDE, 2006). Considérant l'importance du rôle qu'ils ont à jouer dans le succès et l'amélioration des écoles, il semble pertinent de s'interroger sur leur perception d'efficacité personnelle au travail. Cette perception représente la croyance que l'individu a en ses capacités à produire des effets désirés (Bandura, 1997). Elle s'avère une ressource personnelle qui tend à influencer différents aspects de sa vie professionnelle, notamment son engagement et sa santé psychologique (Hobfoll, 2002).

Devant l'absence d'outils psychométriques reconnus, la présente étude vise le développement et la validation d'une échelle qui évalue la perception d'efficacité personnelle des directeurs d'école, et ce, en fonction de leurs rôles professionnels. Pour ce faire, une étude menée auprès de 565 directions d'école (334 femmes et 231 hommes; M âge = 45 ans) a été réalisée. Les résultats soutiennent la structure factorielle de l'échelle et la cohérence interne des items. De plus, ils appuient sa validité de construit par une série de corrélations entre ses sous-échelles et des variables théoriquement associées (p. ex., engagement et épuisement professionnel). Les résultats apportent donc un appui préliminaire à la validité et à la fidélité de l'échelle. Ils mettent également en évidence l'importance de la spécificité des rôles professionnels lors de l'évaluation de la perception d'efficacité personnelle des directeurs d'école.



## Vendredi 27 novembre 2009

8 h 30 – 9 h 00	Inscription
9 h 00 – 10 h 15	Conférence d'ouverture
10 h 15 – 10 h 30	Pause
10 h 30 – 11 h 00	Communications Bloc F
11 h 15 – 11 h 45	Communications Bloc G
11 h 45 – 13 h 30	Dîner de l'ADMEE
13 h 30 – 14 h 00	Communications Bloc H
14 h 15 – 14 h 45	Communications Bloc I
15 h 00 – 15 h 30	Communications Bloc J

### CONFÉRENCE D'OUVERTURE (VENDREDI de 9 h à 10 h 15)

#### Devrait-on se fonder sur les évaluations à grande échelle pour orienter les politiques en éducation?

*Conférencier* : Pierre Brochu, Conseil des ministres de l'éducation du Canada

Depuis quelques années, la prise de décisions fondée sur les faits (*evidence-based decision making*) est populaire auprès des responsables de l'élaboration des politiques en éducation. Lorsque vient le moment d'élaborer de nouvelles politiques ou d'en modifier d'existantes, les systèmes éducatifs ont besoin de données. Les évaluations à grande échelle sont particulièrement attrayantes car elles possèdent les qualités intrinsèques recherchées telles que grande validité, stabilité et comparabilité au fil du temps, en plus d'être assorties d'un éventail de questionnaires contextuels offrant des explications du rendement des élèves dans les matières de base que sont la lecture, les mathématiques et les sciences.

Analyses de corrélation simples ou multiples, covariance, régression, modélisation multi-niveaux, analyses factorielles et autres sont utilisées dans le but de trouver les facteurs qui seront utiles aux politiques et qui pourraient servir pour améliorer le rendement (par exemple, effectif des classes, classes du même sexe, nombre d'heures d'enseignement, imposition des devoirs, ou formation du personnel enseignant).

Cette présentation montrera les possibilités en matière d'analyse qu'offrent les évaluations à grande échelle administrées au Canada pour éclairer les politiques servant aux évaluations. On y discutera également des limites de

ces évaluations dans leur « pouvoir explicatif » et des pratiques exemplaires dans le domaine.

### BLOC F - (VENDREDI de 10 h 30 à 11 h 00)

#### F1 Évolution de l'évaluation des apprentissages au Québec du début du XX<sup>e</sup> siècle à nos jours : politiques et pratiques évaluatives

*Conférenciers et conférencière* : Adrien Cantat, Pierre Valois (Université Laval) et Marie-Hélène Hébert (UQAR)

Depuis la réforme de 2001, l'approche par compétences est à l'heure actuelle la clé de voûte du système éducatif québécois jetant les bases d'un type d'évaluation jusque-là inédit. En prenant appui sur *Le programme des écoles primaires et complémentaires* (Comité catholique, 1923), *La politique générale d'évaluation pédagogique* (MEQ, 1981) et *La politique d'évaluation des apprentissages* (MEQ, 2003), notre exposé mettra en lumière l'évolution de l'évaluation en contexte scolaire au Québec. Pour ce faire, nous étudierons, dans un premier temps, les discours évaluatifs propres à chacun de ces documents en mettant en lumière leurs principales caractéristiques, puis, dans un deuxième temps, nous analyserons les pratiques concrètes d'évaluation utilisées dans les classes durant la mise en application de ces politiques.

#### F2 Expertest, une percée dans le milieu des entreprises et des corporations

*Conférenciers* : Yves Groulx et Jacques Bellemare (Société GRICS)

La Direction des services pédagogiques et de l'évaluation de la société GRICS (Gestion du réseau informatique des commissions scolaires) développe depuis quelques années des produits et des services en évaluation pour une nouvelle clientèle, essentiellement composée d'entreprises et de corporations locales et internationales.

Même si les pratiques évaluatives sont différentes entre une école et une entreprise, les concepts en évaluation restent néanmoins les mêmes. Tout comme l'école, un organisme doit qualifier, certifier ou évaluer des personnes, et doit donc posséder des outils valides et fidèles pour accomplir ces tâches. Les impacts liés aux outils d'évaluation peuvent être considérables sur la crédibilité d'un organisme et sur l'avenir d'une personne évaluée.



Dans cet atelier, nous présenterons succinctement les produits et services qu'offre la Direction des services pédagogiques et de l'évaluation pour des organismes autres que les commissions scolaires. Ensuite, nous présenterons un regard réflexif sur les deux premières années d'exploitation de ce nouveau service. Finalement, nous aborderons les enjeux et les défis qui nous attendent dans ce nouveau marché en pleine expansion.

### **F3 L'évaluation authentique en contexte d'éducation physique au primaire**

*Conférencière et conférenciers* : Johanne Grenier, Gilles Raïche (UQAM), Pierre Boudreau (Université d'Ottawa) et Jean-François Richard (Université de Moncton)

L'évaluation des apprentissages en éducation physique est depuis longtemps un sujet épineux qui ne laisse personne indifférent. Avec la venue des programmes élaborés par compétences, les pratiques évaluatives s'inspirent de l'évaluation authentique (Wiggins, 1989). En s'appuyant sur les cinq dimensions du modèle multidimensionnel de Gulikers et coll. (2004) nous avons élaboré un questionnaire autoadministré sur les pratiques évaluatives déclarées par des éducateurs physiques du primaire. Il semble que les éducateurs physiques (N = 323) présentent des pratiques évaluatives très conformes à l'évaluation authentique pour la dimension « critères et standards » alors qu'ils s'en éloignent beaucoup pour la dimension « contexte physique ». De plus, l'analyse statistique de notre questionnaire démontre que seulement deux dimensions sont révélatrices de l'authenticité de l'évaluation. Ces résultats mettent en doute les dimensions pressenties dans le questionnaire et conduisent à l'élaboration d'un modèle adapté pour l'évaluation authentique en contexte d'évaluation des apprentissages au primaire. Nous présenterons le premier questionnaire, un aperçu des résultats, le deuxième questionnaire qui a été produit à la suite de la validation statistique du premier, ainsi qu'un modèle adapté d'évaluation authentique présentant trois dimensions.

## **BLOC G - (VENDREDI de 11 h 15 à 11 h 45)**

### **G1 L'intégration des pratiques d'évaluation des apprentissages aux pratiques pédagogiques dans le contexte des approches par compétences**

*Conférencières et conférencier* : Diane Leduc, Hélène Meunier, Nadine Talbot et Gilles Raïche (UQAM)

La dernière politique d'évaluation des apprentissages du ministère de l'Éducation du Québec (Gouvernement du Québec, 2003) reflète clairement les tendances académiques modernes. Ainsi, les fonctions et le processus relatifs à l'évaluation qu'on y énonce soulignent autant l'importance du passage du paradigme de l'enseignement à celui de l'apprentissage que les deux fonctions principales de l'évaluation, soit l'aide à l'apprentissage et la reconnaissance des compétences. Qui plus est, la politique met en évidence l'importance de la régulation de la démarche de l'élève et de l'enseignant ainsi que l'évaluation en vue d'une contribution à la réussite éducative.

Dans ce contexte, et dans le cadre d'un projet de recherche subventionné par le FQRSC, nous désirons obtenir un portrait général de la situation à tous les ordres d'enseignement afin d'analyser la structure et l'évolution de l'intégration des pratiques d'évaluation des apprentissages aux pratiques pédagogiques dans le contexte des approches par compétence. À cet effet, nous avons construit un questionnaire portant sur les pratiques évaluatives des enseignants. La communication a pour objectif premier d'expliquer, d'une part, les critères d'évaluation retenus pour le questionnaire et, d'autre part, les principes sur lesquels il s'appuie. Dans un deuxième temps, nous présenterons, les différentes versions du questionnaire qui seront validées au cours de la prochaine année scolaire, auprès des étudiants et des enseignants des ordres d'enseignement primaire, secondaire, collégial et universitaire.

### **G2 Les politiques d'évaluation de l'écriture et leur apport quant aux pratiques éducatives**

*Conférencier et conférencière* : Serge Sévigny (Université Laval) et Isabelle Beaudoin (UQAR)

Très peu d'études ont confirmé ou nié l'utilité des grilles descriptives servant à l'évaluation de l'écriture dans le cadre de l'approche par compétence. Ces grilles peuvent-elles être réinvesties en salle de classe

dans le but d'aider les élèves à améliorer certaines de leurs habiletés en écriture? La présente étude s'intéresse à cette problématique en vérifiant si l'habileté en orthographe d'élèves anglophones et francophones diffère selon les scores rendus par les échelles de niveaux de compétence semblables à celles utilisées aux deux cycles du secondaire. Les données proviennent d'un échantillon de 350 productions écrites par des garçons et filles de 13 et 16 ans. Les résultats confirment que les habiletés en orthographe ne varient pas à tous les niveaux de la grille et que certains résultats varient d'un sexe à l'autre. Chez les anglophones, les résultats varient également selon l'âge. La discussion des résultats porte sur la clarté des critères composant une grille descriptive globale, l'interprétation et l'utilité des scores rendus par ce type de grille lorsqu'il est question d'enseignement et d'apprentissage ainsi que le besoin de recherches destinées à explorer davantage les méthodes actuelles et futures d'évaluation de l'écriture en salle de classe.

### G3 Identification des différences sur le plan des composantes de la violence selon la position sur le jeu au soccer, le niveau de jeu et le rôle du joueur

*Conférenciers* : Julien D'Amours-Raymond, Éric Frenette (Université Laval) et Martin Gendron (UQAR)

Le soccer est devenu le sport le plus populaire auprès des jeunes Québécois de 5 à 14 ans, ayant atteint un taux de participation de 44 % en 2005 (Sport Canada, 2008). Le soccer, ainsi que d'autres sports de compétition (Bredemeier & Shields, 2001), ne semble pas pouvoir échapper à l'émergence des problématiques liées à la violence et ses différentes formes (Bufacci, 2005; Debardieux, 2006). Selon Sport Québec (2007, p.1), « [...] la violence dans le sport peut être définie comme étant une agression physique, verbale ou psychologique ayant pour but de blesser, de faire mal ou d'intimider une personne dans une situation où ce comportement n'a aucun lien avec les règlements et les objectifs de compétition du sport. » De plus, très peu d'études se sont intéressées aux rôles possibles (témoin, victime et agresseur) des joueurs face à la violence. L'objectif de la présente étude consiste à identifier les différences sur le plan des composantes de la violence (intimidation verbale, intimidation physique et violence physique; Coulomb & Pfister, 1998; Shields, 1999) selon la position sur le jeu (milieu de terrain, attaquant, défenseur), le niveau de jeu

et le rôle du joueur (témoin, victime, agresseur). Pour ce faire nous aurons recours à trois méthodes de fonctionnement différentiel d'item (FDI) (Manthel-Haenszel, TESTGRAF, régression logistique). Le FDI représente la différence entre la probabilité de mentionner un comportement pour des joueurs rapportant un même niveau de violence, mais appartenant à des groupes distincts (Clauser & Mazor, 1998). Un total de 1 641 joueurs de soccer (859 garçons et 777 filles) de différents niveaux ont répondu au *Questionnaire d'enquête sur le climat dans le soccer amateur au Québec* (Gendron, Debardieux, Bodin, & Frenette, 2006). Les résultats obtenus tendent à démontrer un continuum selon les composantes de violence, la position du joueur et le rôle du joueur.

## BLOC H - (VENDREDI de 13 h 30 à 14 h)

### ~~III La différence entre la correction de textes manuscrits et de textes dactylographiés, puis imprimés à l'ordinateur~~

~~Conférencière et conférenciers : Caroline Godin, Jean-Guy Blais et Michel Laurier (Université de Montréal)~~ **ANNULÉE**

De nos jours, la remise de travaux scolaires dactylographiés côtoie très fréquemment la remise de travaux manuscrits et différentes recherches américaines ont illustré qu'il pouvait y avoir une différence entre la correction des textes manuscrits et celle des textes dactylographiés. Pour explorer cette question dans le cadre d'une épreuve d'écriture de français, un échantillon de 30 textes manuscrits déjà corrigés par une équipe de correcteurs du ministère de l'Éducation, des Loisirs et du Sport du Québec (MELS), provenant d'un prototype d'épreuve ministérielle, a été utilisé. Ces textes manuscrits ont d'abord été dactylographiés tels que produits par les élèves à l'aide d'un logiciel de traitement de texte, puis imprimés et soumis à une nouvelle correction à la même équipe de correcteurs. Les traces de correction sur les textes manuscrits et sur les textes dactylographiés ont été observées et les analyses ont montré que, contrairement à nos attentes, les différences ne se situeraient pas dans le nombre d'erreurs détectées, mais dans le type d'erreurs détectées. Nous avons en effet remarqué une différence dans la correction des erreurs découlant d'un mauvais usage des signes de ponctuation ou découlant de la syntaxe, selon que le texte corrigé était manuscrit ou dactylographié. L'usage de ponctuation dans les titres, notamment, était plus fréquemment

toléré dans les manuscrits que dans les textes dactylographiés. Les erreurs de majuscules en début de phrase ou de ponctuation en fin de phrase ont été détectées plus souvent dans les textes dactylographiés que dans les manuscrits. Enfin, la construction des phrases (syntaxe) semble avoir été corrigée avec plus de rigueur dans les manuscrits que dans les textes dactylographiés.

## H2 L'évaluation des politiques institutionnelles d'évaluation des apprentissages (PIEA) des collèges du Québec

*Conférencière* : Katie Bérubé (Commission d'évaluation de l'enseignement collégial)

Le *Règlement sur le régime des études collégiales* (RREC) donne l'obligation aux collèges du Québec d'adopter et d'appliquer une politique institutionnelle d'évaluation des apprentissages (PIEA). Créée dans la foulée du renouveau pédagogique du collégial en 1993, la Commission d'évaluation d'enseignement collégial (CEEC) s'est donné comme premier mandat d'évaluer les PIEA des collèges. Après avoir mené différentes opérations d'évaluation relatives aux programmes, aux plans de réussite, aux plans stratégiques et aux politiques institutionnelles d'évaluation des programmes (PIEP), la Commission procède maintenant à l'évaluation de l'application des PIEA.

Après une brève présentation du rôle et des mandats de la Commission, cette communication abordera le cadre général et la démarche empruntée par la Commission pour ses travaux d'évaluation entourant les politiques institutionnelles d'évaluation des apprentissages.

## H3 Le jugement crédible en évaluation de programme

*Conférencière et conférencier* : Marthe Hurteau (UQAM) et Pierre Valois (Université Laval)

Toute démarche évaluative devrait générer un jugement. Pour être recevable, encore doit-il être crédible.

La présentation explore les concepts de jugement et de crédibilité en adoptant une nouvelle perspective. Elle s'appuie sur une recension des écrits ainsi que sur les résultats d'une recherche menée sur les conditions requises pour favoriser l'émergence d'un jugement crédible. Finalement, elle exposera les perspectives retenues pour les futurs travaux ainsi que sur les orientations d'un éventuel livre.

## BLOC I - (VENDREDI de 14 h 15 à 14 h 45)

### I1 Autoévaluation d'une épreuve adaptative à deux niveaux pour l'évaluation du rendement scolaire des élèves dans le cadre d'enquêtes à grande échelle

*Conférencières* : Julie Grondin et Nancy Boiteau (UQAR)

Afin de recueillir les données nécessaires à la planification de leurs politiques éducatives, les systèmes éducatifs ont de plus en plus recours à des enquêtes à grande échelle pour l'évaluation du rendement scolaire des élèves. Le *testing* adaptatif par ordinateur est donc devenu un outil d'évaluation qui, depuis quelques dizaines d'années, est grandement utilisé. Il permet en effet d'évaluer rapidement et précisément de larges échantillons d'élèves. Or, de telles évaluations impliquent bien souvent le recours à l'adaptation d'épreuves afin de rejoindre les divers groupes linguistiques ou culturels présents dans la ou les populations visées par l'évaluation. Dans le but d'assurer que ces évaluations soient équitables pour toutes les personnes évaluées, divers organismes renommés ont rédigé des normes pour guider le développement, la traduction et l'adaptation de ces outils. Pour compléter ces normes, certains auteurs ont essayé d'identifier les aspects principaux d'une épreuve adaptative et, au regard des paramètres qu'ils ont pu relever, proposé un modèle d'évaluation pour ce type d'outil. Dans le cadre de nos recherches, nous travaillons sur le développement d'une épreuve bilingue à deux niveaux, un cas particulier de la modalité à niveaux multiples (une forme d'épreuve adaptative) pour l'évaluation du rendement scolaire des élèves. Nous nous proposons donc de faire une autoévaluation de notre outil au regard des paramètres proposés dans ce modèle.

### I2 L'utilisation des politiques institutionnelles d'évaluation des apprentissages pour assurer la validité des inférences d'évaluation des apprentissages

*Conférencière* : Karine Paquette-Côté (UQAM)

Les politiques institutionnelles d'évaluation des apprentissages (PIEA) des collèges permettent-elles de soutenir la production d'inférences d'évaluation valides? Cette communication présente les résultats d'une recherche visant (1) à valider la structure d'argumentation interprétative de Kane (2006) par l'application de cette structure à l'analyse de PIEA du

réseau collégial québécois et (2) à identifier des moyens que peuvent mettre en place les institutions pour chercher à assurer la validité des inférences d'évaluation au regard des apprentissages des étudiants. En réponse au deuxième objectif, cette recherche a permis d'identifier des lignes directrices permettant de chercher à assurer l'argumentation de la validité des inférences d'évaluation en évaluation des apprentissages au collégial. La nature des moyens identifiés dans le contenu des PIEA analysées amène toutefois à conclure qu'un développement dans les PIEA de moyens plus spécifiques en lien avec les lignes directrices identifiées serait souhaitable pour que ces politiques puissent répondre à leur objectif « d'orienter, d'encadrer et de soutenir les activités reliées à l'évaluation des apprentissages » (Conseil des collègues, 1992).

### I3 **Évaluation des cours de la M.Éd entièrement en ligne en administration scolaire et en enseignement, apprentissage et évaluation à l'Université d'Ottawa**

*Conférencières et conférencier* : Hélène Meunier (UQAM), Claire Isabelle, Emmanuel Duplâa (Université d'Ottawa) et Hélène Fournier (CNRC)

En vertu de la loi (1965 : C.137), l'Université d'Ottawa est redevable et responsable de la qualité de l'enseignement tant auprès des étudiants que du corps professoral. Depuis 2006, la Faculté d'éducation offre la première maîtrise professionnelle entièrement en ligne en administration scolaire et en enseignement, apprentissage et évaluation au Canada. Afin d'évaluer cette M.Éd en ligne, au printemps 2009, une équipe de professeurs a élaboré un questionnaire. Cet instrument, disponible en ligne, a permis de recueillir les opinions auprès d'étudiants qui ont déjà suivi des cours en ligne, tout particulièrement dans le programme de la Maîtrise en éducation. Le premier objectif de la communication est de présenter les indicateurs retenus et les objectifs visés par le questionnaire. Dans un deuxième temps, nous détaillerons les résultats obtenus. Finalement, nous présenterons différentes pistes envisagées quant aux améliorations à apporter aux cours en ligne afin d'offrir une expérience éducative de qualité aux étudiants.

## **BLOC J- (VENDREDI de 15 h à 15 h 30)**

### J1 **L'apport des correcticiels dans la correction de textes d'élèves du secondaire**

*Conférencières et conférencier* : Marie Hélène Mireault, Jean-Guy Blais et Marie-Claude Boivin (Université de Montréal)

L'utilisation des technologies fait partie des politiques mises de l'avant dans le programme de formation de l'école québécoise et les ordinateurs sont de plus en plus présents dans les écoles. De toute évidence, ces politiques ne peuvent qu'influencer les pratiques d'évaluation des apprentissages en salle de classe et dans les opérations à grande échelle. Parmi les applications qui se développent rapidement à l'heure actuelle, on retrouve l'utilisation de correcticiels pour corriger tant la qualité de l'argumentation que la qualité de la langue dans les rédactions des élèves. Au Québec, la correction automatique des textes dans un contexte d'évaluation de la compétence « Écrire des textes » pourrait sûrement apporter une contribution bénéfique tant dans une perspective diagnostique que certificative. Nous avons donc voulu étudier les avantages de l'application d'un correcticiel sur les textes d'élèves pour corriger la qualité de la langue écrite lors d'une épreuve prototype de français soumise par le ministère de l'Éducation, du Loisir et du Sport. Ces résultats ont été mis en parallèle avec ceux obtenus lors de la correction habituelle effectuée par des correcteurs humains afin d'établir l'existence d'éventuelles différences et de juger de l'utilité de ces outils de correction lors d'épreuves de français. Nous avons ainsi voulu répondre à la question suivante : quelles sont les différences entre la correction humaine et la correction effectuée par un correcticiel francophone lors de la correction de la qualité de la langue dans des textes d'élèves du secondaire? Les résultats ont permis, notamment, d'établir un lien entre le niveau du texte corrigé et l'efficacité d'Antidote. Les points forts d'Antidote sont la détection des erreurs d'orthographe d'usage (plus constante que chez les correcteurs humains) et l'analyse plus détaillée des erreurs. Ses points faibles touchent principalement le domaine de la syntaxe.

## J2 **Évaluation des attitudes en stage au collégial**

*Conférencière et conférencier* : Raymonde Gosselin (Collège Édouard-Montpetit) et Gilles Raïche (UQAM)

Le stage, dans le système d'éducation collégiale au Québec, nécessite une évaluation et une notation. L'enseignant doit donc se prononcer, à la fin du stage, sur les apprentissages effectués, sur le niveau de performance et attribuer une note en pourcentage. En stage, les habiletés visées sont de l'ordre du savoir (connaissances, habiletés cognitives), du savoir-faire (habiletés, gestes, techniques, interventions) et du savoir-être (attitude, valeurs, qualités personnelles, habiletés affectives...). Ces habiletés, qualifiées de performances complexes, font appel à plusieurs dimensions de l'apprentissage en même temps. Comment les enseignants peuvent-ils évaluer des attitudes en stage? Comment recueillir des données sur les attitudes, les interpréter afin de prendre les meilleures décisions? La présente recherche a pour but de vérifier comment les enseignants du collégial pourraient évaluer l'acquisition d'attitudes en stage. Afin de répondre à cette question, notre projet de recherche vise donc l'atteinte de certains objectifs spécifiques : a) Identifier les attitudes à évaluer en stage au collégial dans des programmes techniques de la santé avec stages et clients; b) Identifier, reliés à chaque attitude, les comportements à évaluer en stage au collégial dans des programmes techniques de la santé avec stages et clients. Une méthodologie de nature quantitative a été choisie en cohérence avec la problématique et les objectifs de recherche. Nous avons donc élaboré un questionnaire pour nous permettre d'effectuer la collecte des données à un échantillon composé de 505 enseignants en formation technique de la santé comprenant des stages avec clients, au collégial. Les données, recueillies des 247 participants, nous permettront de présenter les résultats de cette recherche.

## J3 **L'apprentissage universitaire et la réussite dans le travail : une analyse sur le contexte, les résultats, les compétences acquises et le succès dans la carrière**

*Conférencières et conférencier* : Carla Barroso da Costa, Nathalie Loye et Jake Murdoch (Université de Montréal)

Les institutions d'enseignement supérieur jouent un rôle important dans la société moderne puisqu'elles doivent amener leurs étudiants à développer les compétences utiles à la réussite de leur carrière professionnelle. La présente étude cherche à vérifier les relations qui existent entre les contextes d'apprentissage, les résultats des étudiants dans les institutions supérieures, les compétences attendues par les organisations et le succès dans la carrière. L'étude recourt à des variables semblables à celles de l'étude de Vermeulen et Schmidt (2008) par rapport au concept de succès dans la carrière, telles que les facteurs extrinsèques (salaire initial et salaire actuel) et le facteur intrinsèque (satisfaction dans le travail). Toutefois, la présente étude ajoute la satisfaction des diplômés relativement à leur institution d'enseignement et leur perception sur l'importance du diplôme dans leur vie professionnelle et personnelle pour expliquer le succès dans la carrière. Dans un premier temps, notre présentation décrira le processus d'élaboration d'un modèle théorique reliant les différentes variables (contexte, résultats, compétences, succès). Dans un deuxième temps, les résultats d'une modélisation aux équations structurelles (SEM) permettront de valider le modèle théorique ou de le faire évoluer jusqu'à obtenir un modèle satisfaisant. Les données sont issues du projet Reflex (The Flexible Professional in the Knowledge Society: New demands on Higher Education in Europe) qui est une enquête sur le devenir professionnel de diplômés de l'année 2000 dans 13 pays européens et au Japon. La troisième partie de la présentation visera à interpréter le modèle SEM obtenu et à identifier les données qui pourraient aider à l'amélioration des politiques évaluatives en lien avec l'intention majeure de la recherche qui est de distinguer les aspects de l'apprentissage universitaire évalués à l'université qui favorisent la réussite professionnelle.