

# Sélection des candidats en médecine : validité prédictive des mini entretiens multiples en contexte francophone

*Admission process of applicants in medicine: predictive validity of Multiple Mini Interviews in a French-speaking context*

Jean-Sébastien RENAUD<sup>1</sup>, Adrien CANTAT<sup>1</sup>, Sawsen LAKHAL<sup>2</sup>, Martine BOURGET<sup>1</sup>, Christina ST-ONGE<sup>3</sup>

<sup>1</sup> Faculté de médecine, Université Laval, Québec, Canada

<sup>2</sup> Faculté d'éducation, Université de Sherbrooke, Québec, Canada

<sup>3</sup> Faculté de médecine et des sciences de la santé, Université de Sherbrooke, Québec, Canada

Manuscrit reçu le 10 janvier 2016 ; commentaires éditoriaux formulés aux auteurs le 4 avril 2016 ; accepté pour publication le 4 juillet 2016

## Mots-clés

Mini Entretiens  
Multiples ; validité  
prédictive ;  
admission ; sélection  
des candidats

**Résumé – Contexte :** Les mini entretiens multiples (MEM) sont un outil de sélection développé par une équipe de l'Université de McMaster pour évaluer les habiletés non-cognitives des candidats en médecine. La capacité des MEM à prédire les résultats aux études en médecine en contexte francophone reste à démontrer. **But :** L'objectif de cette étude est d'évaluer la validité prédictive des MEM en contexte francophone en utilisant les mini entretiens multiples francophones intégrées (MEMFI), développées conjointement par les trois facultés de médecine francophones du Québec. **Méthodes :** Nous avons utilisé un échantillon de 893 étudiants inscrits au programme de doctorat en médecine de l'Université Laval. Les variables indépendantes étaient les résultats aux MEMFI et les résultats académiques antérieurs. Les variables dépendantes étaient les résultats en médecine 1) aux cours intégrateurs, 2) aux cours systèmes, 3) à l'examen longitudinal annuel (ELA) et 4) à l'externat. Des analyses de régression linéaire « pas à pas » ont été réalisées afin de déterminer la valeur prédictive des deux critères de sélection sur les variables dépendantes. **Résultats :** Les résultats aux MEMFI sont surtout associés à ceux à l'externat ( $\beta = 0,268, p < 0,001$ ). Ils prédisent également de manière statistiquement significative les résultats aux cours intégrateurs ( $\beta = 0,086, p = 0,020$ ) et à l'ELA ( $\beta = 0,104, p = 0,019$ ), mais leur pouvoir prédictif est faible et inférieur à celui des résultats académiques antérieurs. **Conclusion :** Les MEM font preuve de validité prédictive en contexte francophone.

**Keywords**

Multiple Mini Interviews; predictive validity; admission; selection of applicants

**Abstract – Background:** Multiple Mini Interviews (MMI) are an admission process developed at McMaster University to assess non-cognitive abilities of applicants in medicine. The ability of MMIs to predict academic success in medical schools has never been demonstrated in a French-speaking context. **Purpose:** The purpose of this study was to assess the predictive validity of MMIs in a French-speaking context using the Integrated French Multiple Mini Interviews (IF-MMI), developed jointly by the 3 French-speaking medical schools in Quebec. **Methods:** We used a sample of 893 students enrolled in the medical studies program at Laval University. Independent variables were IF-MMI scores and previous academic results. Dependent variables were academic results in medicine, including 1) integrative courses, 2) regular courses, 3) annual progress test, and 4) clerkships. We ran stepwise regression analyses to estimate the ability of IF-MMI scores and previous academic results to predict the 4 dependent variables. **Results:** IF-MMI scores are associated with clerkship scores ( $\beta = 0.268$ ,  $p < 0.001$ ). IF-MMI scores also have a weak but significant relation with results in integrative courses ( $\beta = 0.086$ ,  $p = 0.020$ ) and the annual progress test ( $\beta = 0.104$ ,  $p = 0.019$ ). **Conclusion:** MMIs have predictive validity in a French-speaking context.

## Introduction

La sélection des candidats pour l'admission aux études en médecine est un processus à enjeux élevés. D'une part, c'est l'unique porte d'entrée vers la profession pour la majorité des candidats. D'autre part, il importe de sélectionner les candidats les plus à même de réussir et de répondre aux besoins de la société. Qui plus est, en Amérique du Nord, il est impératif d'être en mesure de défendre l'acceptation et le refus des candidats en cas de contestation. Dans ce contexte, le processus utilisé doit être rigoureux, les outils d'évaluation doivent présenter de bonnes qualités psychométriques et l'utilisation des scores doit être justifiable et valide. Un processus de sélection judicieux doit donc être mis sur pied pour choisir les candidats aux études en médecine. Ce processus doit être basé sur les qualités nécessaires pour assumer le rôle de médecin. Selon le cadre de compétences CanMEDS du Collège royal des médecins et chirurgiens du Canada<sup>[1]</sup>, le rôle du médecin est multidimensionnel. Il est attendu qu'un médecin soit un expert de contenu abouti, qu'il démontre des compétences de gestion, de communication et de collaboration, tout en étant un acteur majeur dans la promotion de la santé, ainsi qu'un professionnel respectueux de l'éthique exerçant de manière érudite. En contexte d'admission, ceci se traduit par la

nécessité d'avoir recours à différents outils de sélection. Chaque nouveau contexte d'utilisation de ces outils requiert de mettre en place un processus de validation rigoureux<sup>[2]</sup>.

### Problématique

Un des critères de sélection les plus connus concerne les aptitudes académiques des candidats, aussi appelées habiletés cognitives ; les résultats académiques antérieurs servant d'indicateurs. C'est un critère de sélection important étant donné sa capacité à prédire le succès académique dans un programme de médecine<sup>[3]</sup>. Il est maintenant reconnu que les aptitudes non-académiques, souvent appelées « habiletés non-cognitives », doivent aussi être prises en considération dans la sélection des futurs médecins<sup>[4]</sup>. Ces dernières peuvent être définies comme des qualités personnelles, différentes des aptitudes académiques, qui concourent au succès des étudiants<sup>[5]</sup>. Ces habiletés regroupent une grande variété de caractéristiques individuelles, telles que les habiletés de communication et de collaboration, le professionnalisme, le sens éthique, les attitudes, les traits de personnalité ou les motivations<sup>[6, 7]</sup>, et sont complémentaires des aptitudes académiques dans la prédiction de la performance future des candidats<sup>[8]</sup>. Toutefois, il faut garder en tête que cette

dichotomie entre les habiletés dites cognitives et non-cognitives possède un caractère artificiel. Il apparaît effectivement peu probable que les habiletés dites non-cognitives n'impliquent aucun processus cognitif, et réciproquement.

Au Canada, une équipe de l'Université McMaster a développé un outil de sélection, désigné par la locution « mini entretiens multiples (MEM) », afin d'évaluer ces habiletés non-cognitives chez les candidats en médecine<sup>[6]</sup>. Cet outil s'apparente, par sa forme, à l'Examen Clinique Objectif Structuré (ECOS), élaboré par Harden dans les années 1970 en Écosse à l'Université de Dundee<sup>[9]</sup>. Les MEM visent à évaluer la performance des candidats à travers une séquence de brèves entretiens ou de mises en situation avec ou sans acteur, présentées dans différentes « stations », constituant un « circuit ». Chaque candidat est ainsi évalué dans plusieurs situations par un examinateur préalablement « standardisé » utilisant une grille d'évaluation spécifique à la situation observée. En bref, l'idée derrière les MEM est la même que pour l'ECOS : diminuer l'erreur de mesure en évaluant la performance de chaque candidat dans plusieurs situations et par plusieurs examinateurs plutôt que dans une seule situation par un ou quelques examinateurs.

Jusqu'à présent, la plupart des études de validation des MEM ont été conduites en contexte anglophone, que ce soit aux États-Unis, en Australie, en Grande-Bretagne, ou au Canada anglais<sup>[10]</sup>. L'effort de validation des MEM en contexte francophone demeure donc à poursuivre. En effet, le contexte francophone diffère du contexte anglophone par sa langue et sa culture, mais aussi par son système d'éducation. Par exemple, au Québec, une province canadienne francophone, ce système a pour particularité la présence d'un cursus dit collégial permettant la transition entre l'école secondaire et l'université. Les futurs étudiants en médecine au Québec ont généralement fait deux années d'études collégiales avant de postuler et par la suite intégrer une faculté de médecine. Par exemple, à l'Université Laval, selon les années, environ 55 % à 60 % des candidats québécois admis proviennent du collégial. En milieu anglophone, les candidats doivent compléter un baccalauréat universitaire de trois ou quatre ans avant de

pouvoir faire une demande d'admission en médecine. Pour cette raison, la moyenne d'âge des candidats en médecine au Québec est inférieure à celle observée en milieu anglophone. À cause de ces différences, il est impossible d'assumer que la validité des MEM sera la même en contexte francophone qu'en contexte anglophone.

Les études réalisées jusqu'à ce jour en contexte anglophone ont démontré que les MEM peuvent servir à évaluer les habiletés non-cognitives des candidats en sciences de la santé de manière fiable et valide, comme le conclut une méta-analyse réalisée sur le sujet<sup>[11]</sup>. Leur fidélité, estimée par le coefficient alpha de Cronbach, varie entre « acceptable » ( $\alpha = 0,59$  à  $0,79$ )<sup>[12]</sup> et « bonne » ( $\alpha = 0,86$ )<sup>[13]</sup>. Quant à leur généralisabilité, les chercheurs rapportent des coefficients  $G$  de  $0,76$ <sup>[14]</sup> à  $0,79$ <sup>[15]</sup>. Des études documentent également la validité des MEM. Cet outil de sélection s'avère être corrélé avec des construits distincts, mais associés, tels que les aptitudes cliniques ( $r = 0,31$ ,  $p < 0,10$ )<sup>[16]</sup>, le raisonnement dans les sciences sociales et humaines et la communication écrite ( $r = 0,20$ ,  $p < 0,01$ )<sup>[11]</sup>. Ils ont également une forte validité d'apparence<sup>[17]</sup> et sont acceptés aussi bien par les évaluateurs que par les candidats<sup>[12]</sup>, sans compter qu'ils ne semblent pas affectés par un effet de « *coaching* »<sup>[18]</sup>. Les MEM prédisent aussi les futures performances académiques des candidats. Eva et al.<sup>[19]</sup> ont identifié une relation positive significative entre les résultats aux MEM et ceux obtenus à l'ECOS administré au pré externat à l'Université McMaster ( $\beta = 0,44$ ,  $p < 0,01$ ). Reiter et al.<sup>[20]</sup>, pour leur part, ont trouvé que les résultats aux MEM sont associés aux notes attribuées par les directeurs de l'externat ( $\beta = 0,70$ ,  $p < 0,001$ ) et aux évaluations réalisées (*encounter card ratings*) par les superviseurs cliniques ( $\beta = 0,50$ ,  $p < 0,01$ ). À l'université de Dundee, Husbands et Dowell<sup>[21]</sup> ont trouvé que les MEM sont le meilleur prédicteur des résultats au pré externat, expliquant 6 % de la variance pour l'examen écrit et 11 % pour l'ECOS lors de la première année d'études et 7 % pour celui de la deuxième année. Enfin, quelques études identifient un lien entre la performance aux MEM et les résultats aux examens certificatifs encadrant la profession médicale. En effet, Eva et al.<sup>[8]</sup> rapportent, suite à des analyses de

régression linéaire multiples, des coefficients standardisés  $\beta$  de 0,12 entre les MEM et le score à la partie I de l'examen d'aptitude du Conseil médical du Canada (EACMC partie I) et de 0,21 avec la partie II (EACMC, Partie II). De plus, Eva et al.<sup>[14]</sup> mentionnent des corrélations de l'ordre de 0,43 ( $p < 0,05$ ) entre les MEM et le nombre de stations ECOS réussies à l'EACMC partie II et de 0,36 ( $p < 0,10$ ) avec le score total à celui-ci, alors que les outils de sélection traditionnels, comme l'entrevue individuelle, n'y sont pas associés de manière significative.

Quelques études ont fait avancer les connaissances sur la validité des MEM en contexte francophone. Celles réalisées jusqu'à maintenant rapportent une cohérence interne ( $\alpha = 0,70$ )<sup>[22]</sup> et une généralisabilité ( $G = 0,745$ )<sup>[23]</sup> satisfaisantes<sup>[24]</sup>. Toutefois, la capacité des MEM à prédire les résultats aux études en médecine en contexte francophone reste à démontrer. La finalité d'un outil de sélection étant d'aider à identifier les candidats les plus susceptibles de réussir, l'évaluation de la validité prédictive des MEM en contexte francophone est primordiale.

## Contexte de l'étude

Devant la nécessité d'évaluer les habiletés non cognitives des candidats aux études en médecine, d'une part, et l'accumulation de résultats de recherches favorables à l'utilisation des MEM comme outil de sélection, d'autre part, les trois facultés de médecine francophones de la province de Québec (les universités Laval (Québec), de Sherbrooke et de Montréal) se sont associées pour élaborer et administrer leurs propres MEM, nommées mini entrevues multiples francophones intégrées (MEMFI). Les MEMFI sont élaborées conjointement par ces trois facultés. De plus, elles sont administrées simultanément dans quatre sites, soit à Québec, Montréal, Sherbrooke et à son campus décentralisé de l'Université de Moncton au Nouveau-Brunswick.

Le programme de doctorat en médecine de l'Université Laval comporte deux parties, le pré externat et l'externat. Le pré externat peut être fait en deux ans, deux ans et demi ou trois ans, au choix de

l'étudiant. Il est composé de cours qualifiés d'intégrateurs et de systèmes. Les cours intégrateurs ont comme objectifs le développement des habiletés de raisonnement clinique et de résolution de problème des étudiants au doctorat en médecine. Plusieurs modalités d'évaluation sont utilisées dans le cadre de ces cours afin d'apprécier les connaissances et compétences des étudiants, soit les travaux en équipe et individuels, les examens écrits, les examens pratiques avec patients simulés et l'ECOS. En parallèle, les étudiants suivent des cours de contenus dits systèmes, portant sur les différents systèmes du corps humain (urinaire, digestif, respiratoire, ...) et intégrant les aspects histologique, physiologique, pathologique et clinique. L'externat, quant à lui, a une durée de 20 mois. Il comprend plusieurs périodes de stages cliniques, allant de trois à six semaines, réalisées par les étudiants dans les établissements du Réseau universitaire intégré en santé de l'Université Laval. Ces stages cliniques correspondent chacun à une spécialité médicale, telle que chirurgie, psychiatrie, pédiatrie, gynécologie, etc. Enfin, depuis la session d'automne 2013, la faculté de médecine de l'Université Laval utilise un examen longitudinal annuel (ELA) qui s'adresse à tous les étudiants inscrits au programme de doctorat en médecine. Celui-ci est un test de rendement progressif, de type formatif et similaire au *progress test* utilisé par l'Université de Maastricht<sup>[25]</sup>, visant à aider les étudiants à intégrer l'ensemble des objectifs d'apprentissage du Conseil médical du Canada tout au long de leur cheminement. En d'autres termes, son but est d'apprécier la construction graduelle et organisée de la base de connaissances nécessaire à la résolution de problèmes complexes en médecine.

## Objectif de l'étude

L'objectif de cette étude est d'évaluer la validité prédictive des MEM en contexte francophone. Plus précisément, l'étude met en relation les résultats académiques des étudiants admis en médecine à l'Université Laval avec leurs résultats aux MEM et leurs résultats au programme de doctorat en médecine. Les résultats académiques antérieurs sont considérés

dans cette étude pour refléter le fait que, dans la réalité, les outils de sélection des habiletés non-cognitives sont utilisés de manière complémentaire plutôt que seuls pour sélectionner les candidats en médecine. Les résultats de l'étude seront donc davantage représentatifs de la réalité.

## Matériels et Méthodes

### Description des mini entretiens multiples francophones intégrés

Les candidats aux MEMFI doivent effectuer un parcours comportant 12 stations de sept minutes, évaluées par 12 examinateurs différents, soit un par station, auxquelles s'ajoutent deux stations de repos. Le choix du nombre de stations et de leur durée a été influencé par une étude pilote et une analyse de généralisabilité dont les résultats sont publiés<sup>[23]</sup>. Des exemples de stations de MEM sont présentés dans l'article d'Eva et al.<sup>[6]</sup>. Chacune des stations porte sur une mise en situation différente. Deux catégories de mises en situation sont employées. La première est de type jeu de rôle et nécessite le recours à un examinateur et à un acteur. Ces stations permettent de simuler une grande variété de situations permettant d'observer les candidats en interaction avec une autre personne. La seconde catégorie de mises en situation, celles de type discussion, nécessite uniquement le recours à un examinateur. Ces stations permettent de mettre les candidats dans des situations où ils ont une tâche à réaliser ou encore de les faire discuter avec un examinateur sur différents sujets. Pour chacune d'elles, la performance des candidats est évaluée à l'aide d'une grille d'évaluation globale élaborée spécialement pour les MEMFI. Cette grille est inspirée de ce que Scallon<sup>[26]</sup> appelle la grille descriptive globale qu'il suggère pour évaluer des productions complexes de manière à inférer des compétences. Elle décrit des niveaux de performance qualitativement distincts auxquels un pointage correspond et la tâche des examinateurs est de juger où se situe la performance de chaque candidat par rapport à ces niveaux de performance. En d'autres

termes, les MEMFI reposent sur le jugement et permettent une construction méthodique de celui-ci en le balisant à l'aide d'une grille d'évaluation décrivant des niveaux qualitatifs distincts de performance et en multipliant les observations.

La passation des MEMFI se fait simultanément sur quatre sites différents. La standardisation des MEMFI, des patients simulés et des évaluateurs se fait de manière concertée entre les trois facultés de médecine impliquées. Lorsque nécessaire, les incidents survenant durant la passation sont gérés conjointement en temps réel par les trois facultés de médecine. Seuls les aspects purement logistiques sont sous la responsabilité de l'équipe chargée de la passation sur chacun des sites. Les scores obtenus aux MEMFI lors de cette passation multi-sites sont valables pour les trois facultés de médecine impliquées. Cependant, chacune de ces facultés possède son propre programme de doctorat en médecine et utilise les scores aux MEMFI dans le cadre du processus d'admission comme elle le souhaite, notamment en ce qui concerne la pondération qui leur est attribuée.

### Échantillon

Nous avons utilisé un échantillon de 893 étudiants inscrits au programme de doctorat en médecine de l'Université Laval. Il était composé de la totalité des étudiants des quatre années d'admission suivantes : 2009 ( $n = 215$ ), 2010 ( $n = 225$ ), 2011 ( $n = 228$ ) et 2012 ( $n = 225$ ). La base de données utilisée pour l'étude ne contenait aucune information susceptible d'identifier les candidats, comme le sexe, l'âge ou l'ethnie. Pour cette raison, il n'est pas possible de décrire davantage les caractéristiques de l'échantillon.

### Variation indépendante

### Résultats aux MEMFI

Les scores bruts obtenus par les candidats à l'issue des MEMFI sont standardisés par site et par jour de passation en les convertissant en scores Z.



Ces derniers sont ensuite exprimés sur une échelle ayant une moyenne de 500 et un écart-type de 50. Dans le cadre de cette étude, les scores des étudiants composant notre échantillon varient entre 400 et 630. De 2009 à 2012, le coefficient de fidélité alpha de Cronbach des MEMFI a varié entre 0,65 et 0,73.

### **Résultats académiques antérieurs**

Au Québec, la performance académique antérieure des étudiants est mesurée à l'aide d'une cote de rendement (CR). Qu'il s'agisse de la cote de rendement au collégial ou à l'université, elle tient compte de la position de l'étudiant par rapport à son groupe et d'un indice de la force relative de groupe.

### Variables dépendantes

La performance académique des étudiants en médecine est mesurée à l'aide de leurs résultats 1) aux cours intégrateurs ; 2) aux cours systèmes ; 3) à l'ELA et : 4) à l'externat. Les habiletés non-cognitives des étudiants, liées à la communication, à la collaboration ou au professionnalisme, par exemple, sont principalement observées à l'externat et dans les cours intégrateurs.

### **Performance aux cours intégrateurs et systèmes**

À l'Université Laval, les scores en pourcentage aux différents cours sont converties en lettres (de E à A<sup>+</sup>) selon un barème propre à chaque cours ou à chaque programme, selon le cas. Ces lettres correspondent quant à elles à une valeur numérique, variant entre 0,00 pour E à 4,33 pour A<sup>+</sup>, servant au calcul de la moyenne cumulative de chaque étudiant. En cohérence avec ce système, nous avons utilisé ces scores sur 4,33 pour calculer la moyenne cumulative des étudiants aux cours intégrateurs et systèmes.

### **Performance à l'examen longitudinal annuel (ELA)**

Cet examen est composé de 180 questions à choix multiples portant sur six domaines : la chirurgie, la gynécologie obstétrique, la médecine, la

pédiatrie, la psychiatrie et la santé publique. Il est administré chaque année au début de la session d'automne et est complété par tous les étudiants de la première à la cinquième année du programme de doctorat en médecine. Des comités d'experts de chacun des sept domaines rédigent les questions. La cohérence interne ( $\alpha = 0,92$ ) est très satisfaisante pour l'édition de 2013, utilisée pour cette étude. Les scores Z des étudiants ont été calculés par cohorte sur la totalité de l'examen et par domaine d'étude à savoir chirurgie, gynécologie, médecine, pédiatrie, psychiatrie, santé publique.

### **Performance à l'externat**

Les périodes de stage sont évaluées à l'aide d'une fiche d'évaluation dans laquelle le superviseur doit se prononcer sur cinq dimensions : l'expertise clinique (10 questions,  $\alpha = 0,81$ ), la communication (8 questions,  $\alpha = 0,81$ ), la collaboration (2 questions,  $\alpha = 0,76$ ), l'érudition (5 questions,  $\alpha = 0,70$ ) et le professionnalisme (7 questions,  $\alpha = 0,84$ ). La performance à chacune de ces dimensions est évaluée sur une échelle en quatre points (supérieure, attendue, limite, insuffisante) avec la possibilité de sélectionner « ne s'applique pas », et la somme des réponses aux items est considérée comme une échelle de niveau à intervalle<sup>[27]</sup>. Une étude préalable a été menée afin d'apprécier les qualités psychométriques de cette grille d'évaluation. La cohérence interne s'avère être très satisfaisante ( $\alpha = 0,92$ ), alors que les analyses factorielles confirmatoires (AFC) réalisées soutiennent le découpage théorique en cinq dimensions de la grille d'évaluation.

### Analyses

Des analyses de régressions linéaires « pas à pas » ont été réalisées afin de déterminer la valeur prédictive des deux critères de sélection (MEMFI et CR) sur la performance académique des étudiants aux cours intégrateurs et systèmes, à l'ELA et à l'externat (résultat global et à chacune des cinq dimensions évaluées). La méthode « pas à pas » a été

**Tableau I.** Régression linéaire prédisant la performance aux cours intégrateurs ( $n = 664$ ) et aux cours systèmes ( $n = 667$ ).

Variables dépendantes	Modèles statistiques				Variables indépendantes			
	Modèles	$R^2$	$F$	$p$	Prédicteurs	B	$\beta$	$p$
Cours intégrateurs	1	0.121	91.517	< 0.001	CR	0.056	0.349	< 0.001
	2	0.129	48.782	< 0.001	CR MEMFI	0.058 0.001	0.361 0.086	< 0.001 0.020
Cours systèmes	-	0.152	119.598	< 0.001	CR	0.094	0.390	< 0.001

$R^2$  : Coefficient de détermination

$F$  : Valeur du test  $F$  qui permet de vérifier le niveau de signification statistique du modèle

$p$  : Probabilité de l'hypothèse nulle

B : Coefficient de régression linéaire non standardisé

$\beta$  : Coefficient de régression linéaire standardisé

CR : Cote de rendement

MEMFI : Mini entretiens multiples francophones intégrés

retenue pour dégager le ou les critères de sélection qui contribuent de manière significative à prédire la performance académique en termes de pourcentage de variance expliquée, exprimé par le coefficient de détermination  $R^2$ . Les analyses ont été réalisées à l'aide du logiciel SPSS version 22.

Les tailles d'échantillons pour ces analyses de régression sont variables. En effet, l'avancement dans le programme varie selon plusieurs facteurs : année d'admission, durée du pré externat, parcours personnel, etc. Ainsi, les analyses de régression visant à prédire la performance aux cours intégrateurs et systèmes sont respectivement basées sur un échantillon de 664 et de 667 étudiants, celles pour prédire la performance à l'externat sur un échantillon de 242 étudiants et celles pour prédire la performance à l'ELA sur un échantillon de 461 étudiants.

## Résultats

### Cours intégrateurs et systèmes

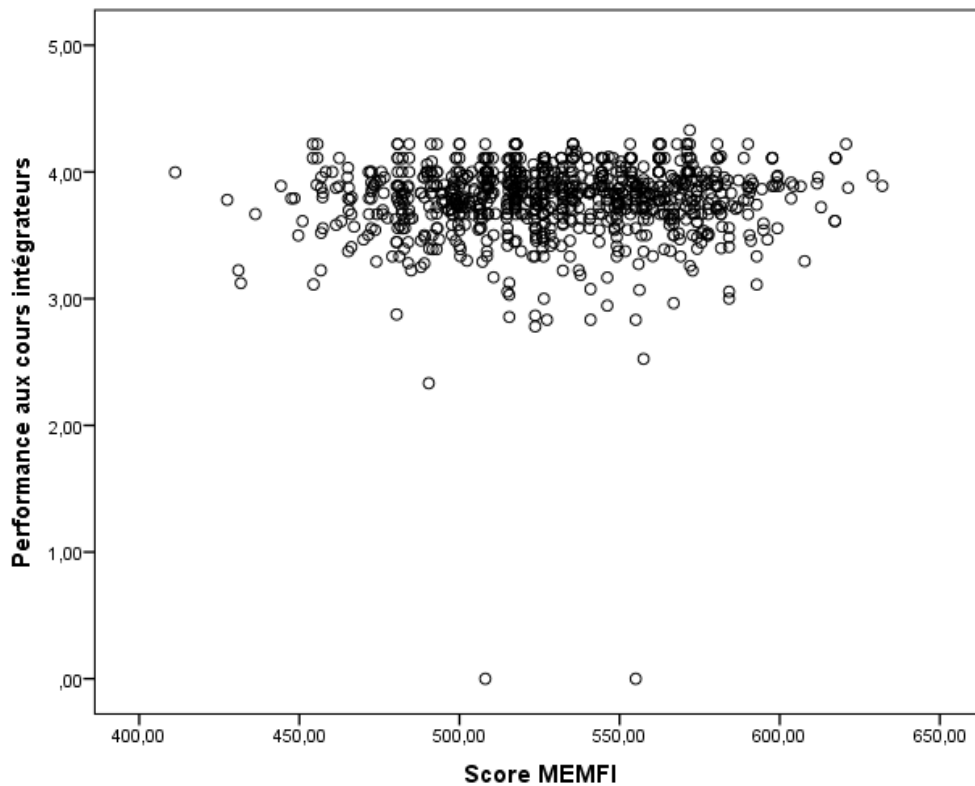
Le tableau I présente les résultats des analyses de régression linéaire visant à déterminer la capacité prédictive des critères de sélection (CR et MEMFI) sur la performance aux cours intégrateurs et non intégrateurs.

Les analyses de régression linéaire se sont toutes révélées statistiquement significatives. Pour prédire la performance aux cours intégrateurs, la méthode « pas à pas » a permis de retenir la CR et les scores aux MEMFI comme prédicteurs significatifs. La CR explique 12,1 % de la variance pour la performance aux cours intégrateurs (modèle 1) alors que la combinaison de la CR et des scores obtenus aux MEMFI en explique 12,9 % (modèle 2). La figure I illustre la relation entre les scores aux MEMFI et ceux aux cours intégrateurs. En ce qui concerne la performance aux cours systèmes, seule la CR apparaît être un prédicteur significatif. Celle-ci explique 15,2 % de la variance de la performance aux cours systèmes. Ces résultats indiquent que la CR est le meilleur prédicteur de la performance académique aux cours, et ce, qu'ils soient de nature intégratrice ou non. Il est à noter que les scores obtenus aux MEMFI n'expliquent que peu voire pas du tout cette performance.

### Examen longitudinal annuel

Les résultats des analyses de régression linéaire visant à établir la validité prédictive de la CR et des scores obtenus aux MEMFI vis-à-vis de la performance à l'ELA sont répertoriés dans le tableau II.

Les analyses de régression sont toutes significatives. Relativement à la prédiction du score Z total à



**Fig. 1.** Nuage de points illustrant l'association entre le score aux mini entretiens multiples francophones intégrées (MEMFI) et la performance aux cours intégrateurs ( $n = 664$ ).

l'ELA, la CR permet d'expliquer 10,9 % de sa variance (Modèle 1), tandis que la combinaison de la CR et du score obtenu aux MEMFI en explique 11,9 % (Modèle 2). Les analyses de régression visant à prédire la performance à chacun des champs disciplinaires de l'ELA ont toutes retenu la CR comme premier prédicteur significatif. Le pourcentage de variance expliquée par la CR dans la performance à ces six champs disciplinaires varie entre 2,2 % (gynécologie) et 7,4 % (chirurgie). Les scores obtenus aux MEMFI apparaissent comme deuxième prédicteur statistiquement significatif, derrière la CR, pour deux des six champs disciplinaires, soit médecine et pédiatrie. Pour ces derniers, l'ajout des scores aux MEMFI à la CR ne permet d'expliquer, respectivement, que 0,8 % et 0,9 % de variance supplémentaire. D'après ces analyses, la CR est le principal prédicteur des scores obtenus à l'ELA, que ce soit pour le total ou pour chacun des champs disciplinaires

composant cet examen. La figure 2 illustre la relation entre le score aux MEMFI et le score total à l'ELA.

#### Externat

Le tableau III présente les résultats issus des analyses de régression linéaire visant à déterminer la valeur prédictive des deux critères de sélection à l'étude sur la performance à l'externat (score total et score à chacune des dimensions évaluées).

Les analyses de régression se sont toutes révélées statistiquement significatives. Les résultats obtenus aux MEMFI expliquent 6,2 % (modèle 1) de la variance au score total à l'Externat, alors que l'association de ce critère et de la CR en explique 12,3 % (modèle 2). Les MEMFI ont été retenues comme les premiers prédicteurs pour quatre des cinq dimensions évaluées lors de l'Externat, suivi par la CR lors du deuxième « pas ». Pour ces quatre dimensions, les



**Tableau II.** Régression linéaire prédisant la performance à l'ELA ( $n = 461$ ).

Variables dépendantes (scores Z)	Modèles statistiques				Variables indépendantes			
	Modèles	$R^2$	$F$	$p$	Prédicteurs	$B$	$\beta$	$p$
Total	1	0.109	56.069	< 0.001	CR	0.190	0.330	< 0.001
	2	0.119	31.066	< 0.001	CR MEMFI	0.197 0.003	0.343 0.104	< 0.001 0.019
Chirurgie	-	0.074	36.512	< 0.001	CR	0.158	0.271	< 0.001
Gynécologie	-	0.022	10.089	0.002	CR	0.085	0.147	0.002
Médecine	1	0.073	35.883	< 0.001	CR	0.156	0.269	< 0.001
	2	0.081	20.245	< 0.001	CR MEMFI	0.163 0.003	0.281 0.094	< 0.001 0.038
Pédiatrie	1	0.029	13.548	< 0.001	CR	0.098	0.169	< 0.001
	2	0.038	9.055	< 0.001	CR MEMFI	0.105 0.003	0.181 0.098	< 0.001 0.035
Psychiatrie	-	0.033	15.752	< 0.001	CR	0.104	0.182	< 0.001
Santé publique	-	0.052	24.992	< 0.001	CR	0.129	0.227	< 0.001

$R^2$  : Coefficient de détermination

$F$  : Valeur du test  $F$  qui permet de vérifier le niveau de signification statistique du modèle

$p$  : Probabilité de l'hypothèse nulle

$B$  : Coefficient de régression linéaire non standardisé

$\beta$  : Coefficient de régression linéaire standardisé

MEMFI : Mini entretiens multiples francophones intégrées

CR : Cote de rendement

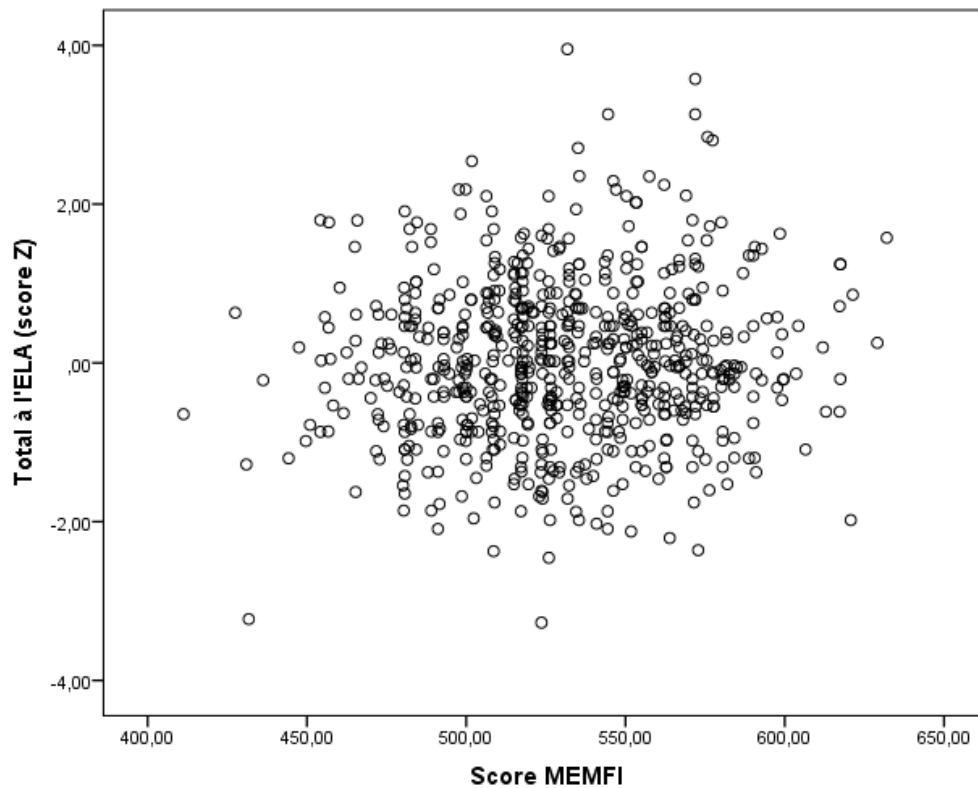
scores aux MEMFI expliquent entre 4,4 % (érudition) et 7,0 % (expertise clinique) de la variance. Seule la dimension « collaboration » se démarque puisque, pour celle-ci, la CR apparaît comme premier prédicteur, expliquant 4,2 % de la variance. La combinaison des scores aux MEMFI et de la CR explique quant à elle entre 5,9 % et 13,0 % de la variance aux cinq dimensions de performance à l'externat. La relation entre le score aux MEMFI et le score total à l'externat est illustrée sur la figure III.

## Discussion

Cette étude visait à évaluer la validité prédictive des MEM en contexte francophone, une étape essentielle de la validation d'un outil de sélection. Les analyses de régression linéaire réalisées ont permis de vérifier si les résultats aux MEMFI sont associés aux résultats en médecine des candidats admis. À la

lumière des résultats obtenus, les MEMFI apparaissent être un outil de sélection faisant preuve de validité prédictive, étant associé de manière significative aux résultats à l'externat, aux cours intégrateurs et à l'ELA.

Les MEMFI présentent une complémentarité avec les résultats académiques antérieurs dans la prédiction rendement des candidats admis en médecine. Les résultats aux MEMFI sont davantage associés à ceux à l'externat, où sont principalement évaluées les habiletés non-cognitives dans le cursus médical à l'Université Laval. Ils expliquent autant de variance dans le score total à l'externat que les résultats académiques antérieurs. Ainsi, en ajoutant les MEMFI aux résultats académiques antérieurs dans la sélection des candidats, ceci a fait doubler le pourcentage de variance expliquée du score total à l'externat, qui se situe maintenant à 12 %. Les MEMFI prédisent également de manière statistiquement significative les résultats aux cours intégrateurs et à l'ELA, mais



**Fig. 2.** Nuage de points illustrant l'association entre le score aux Mini entretiens multiples francophones intégrées (MEMFI) et celui à l'examen longitudinal annuel (ELA) ( $n = 461$ ).

leur pouvoir prédictif est faible et inférieur à celui des résultats académiques antérieurs. Les résultats académiques antérieurs, quant à eux, sont davantage associés aux évaluations qui ciblent les habiletés cognitives, soit les cours intégrateurs et systèmes ainsi que l'ELA. Ces résultats rejoignent ceux obtenus en milieu anglophone, où les MEM étaient davantage associés aux résultats à l'externat qu'aux résultats au préclinique<sup>[20]</sup>. Ainsi, comme l'ont souligné Reiter et al.<sup>[20]</sup> et Eva et al.<sup>[8]</sup>, l'association d'un critère de sélection cognitif et d'une modalité d'évaluation des attributs non-cognitifs en vue de sélectionner de futurs médecins s'avère être plus avantageuse que d'utiliser uniquement le rendement académique antérieur.

Plus spécifiquement, même si le pourcentage de variance expliquée est somme toute modeste, les MEMFI sont le principal prédicteur des résultats des étudiants sur quatre des cinq dimensions évaluées à

l'externat, à savoir l'expertise clinique, la communication, l'érudition et le professionnalisme. Ces résultats permettent de penser que les MEMFI évaluent bien des ressources qui seront nécessaires à la construction ultérieure d'habiletés qui devront être mobilisées et démontrées lors de l'externat, deux à trois années après l'admission en médecine. Néanmoins, la cinquième dimension évaluée à l'externat, la collaboration, est davantage expliquée par les résultats académiques antérieurs. Ce dernier résultat est étrange à première vue et semble difficile à expliquer. Une hypothèse est que la collaboration n'est mesurée que par deux items dans la fiche d'évaluation des externes, contrairement aux autres dimensions qui sont couvertes par cinq à 10 items chacune. En d'autres termes, sa validité de contenu pourrait être moindre, de sorte que cette dimension est moins bien évaluée que les autres. Pour en être plus certain, il faudrait vérifier si l'ajout d'items évaluant la

**Tableau III.** Régression linéaire prédisant la performance à l'externat ( $n = 242$ ).

Variables dépendantes	Modèles statistiques				Variables indépendantes			
	Modèles	$R^2$	$F$	$p$	Prédicteurs	$B$	$\beta$	$p$
Score total	1	0.062	15.828	< 0.001	MEMFI	0.001	0.249	< 0.001
	2	0.123	16.641	< 0.001	MEMFI CR	0.001 0.016	0.268 0.247	< 0.001
Expertise clinique	1	0.070	17.903	< 0.001	MEMFI	0.001	0.264	< 0.001
	2	0.130	17.746	< 0.001	MEMFI CR	0.001 0.018	0.282 0.246	< 0.001
Communication	1	0.053	13.373	< 0.001	MEMFI	0.001	0.230	< 0.001
	2	0.078	10.115	< 0.001	MEMFI CR	0.001 0.013	0.242 0.160	< 0.001 0.011
Collaboration	1	0.042	10.403	0.001	CR	0.023	0.204	0.001
	2	0.059	7.746	0.001	CR MEMFI	0.024 0.001	0.214 0.132	0.001 0.037
Érudition	1	0.044	11.060	0.001	MEMFI	0.001	0.210	0.001
	2	0.076	9.758	< 0.001	MEMFI CR	0.001 0.011	0.224 0.178	< 0.001 0.005
Professionnalisme	1	0.054	13.612	< 0.001	MEMFI	0.001	0.232	< 0.001
	2	0.092	12.005	< 0.001	MEMFI CR	0.001 0.015	0.247 0.195	< 0.001 0.002

$R^2$  : Coefficient de détermination

$F$  : Valeur du test  $F$  qui permet de vérifier le niveau de signification statistique du modèle

$p$  : Probabilité de l'hypothèse nulle

$B$  : Coefficient de régression linéaire non standardisé

$\beta$  : Coefficient de régression linéaire standardisé

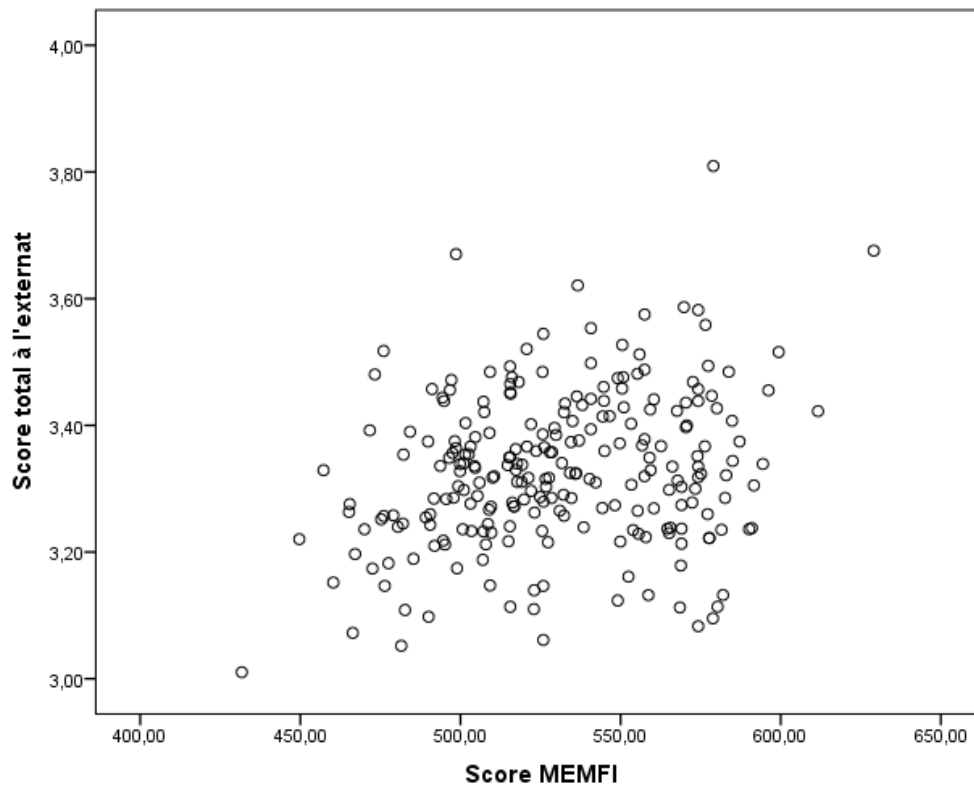
MEMFI : Mini entretiens multiples francophones intégrées

CR : Cote de rendement

collaboration dans la fiche d'évaluation utilisée à l'externat permet d'obtenir une meilleure corrélation avec les résultats aux MEMFI. Une autre hypothèse pour expliquer ce résultat est que ces deux items évaluent les étudiants à l'externat sur leur collaboration avec l'équipe de soins. En revanche, les MEMFI évaluent la collaboration des candidats avec un acteur plutôt qu'avec une équipe, ce qui est une dynamique interpersonnelle différente.

À l'égard de la force de la relation entre les résultats aux MEMFI et ceux obtenus ultérieurement en médecine, elle est ici moindre que celle observée en contexte anglophone. Les MEMFI permettent d'expliquer 6 % de la variance des résultats à l'externat et au maximum 1 % de la variance des résultats

au préclinique. En comparaison, les MEM apparaissent être le meilleur prédicteur au pré externat à l'Université McMaster<sup>[19]</sup> et à l'Université de Dundee<sup>[21]</sup>. Plus précisément, Eva et al.<sup>[19]</sup> ont trouvé que les MEM expliquaient 10 % de la variance des résultats à un ECOS au préclinique. Husbands et Dowell<sup>[21]</sup> ont trouvé que les MEM prédisaient entre 3 % et 12 % des résultats au préexternat. Cette différence pourrait s'expliquer par le fait que les scores établis pour le préexternat dans notre étude se basent sur les moyennes obtenues lors de l'ensemble des cours intégrateurs et systèmes, qui reflètent les résultats à plusieurs cours et plusieurs modalités d'évaluation, dont plusieurs examens écrits. En revanche, pour l'Université McMaster, le score utilisé ne se



**Fig. 3.** Nuage de points illustrant l'association entre le score aux MEMFI et celui à l'externat ( $n = 242$ ). Mini entrevues multiples francophones intégrées (MEMFI).

base que sur le score obtenu à deux ECOS. Pour l'Université de Dundee, la performance au pré externat n'est établie que par un ECOS et un examen écrit (QCM et questions d'association). Par ailleurs, Reiter et al.<sup>[20]</sup> rapportent que les MEM expliquent entre 26 % et 32 % de la variance aux résultats à l'externat. Ces pourcentages de variances expliquées sont élevés par rapport à ceux observés ici et à ceux rapportés par Eva et al.<sup>[14]</sup>, qui avaient trouvé que les MEM expliquaient entre 12 % et 13 % de la variance des scores totaux à un examen de certification (EACMC partie II). Cette différence relativement importante peut s'expliquer de plusieurs manières. D'abord, les MEM varient d'une université à une autre, de sorte que les habiletés non-cognitives évaluées ne sont pas nécessairement les mêmes, ce qui peut affecter leur validité prédictive. Ensuite, il est possible que les outils d'évaluation utilisés à l'externat dans l'étude de Reiter et al.<sup>[20]</sup> portent sur des

éléments plus proches de ceux évalués dans les MEM que dans la présente étude. Par ailleurs, la fiche d'évaluation des externes utilisée dans leur étude est possiblement plus sensible aux différences individuelles, car les items y sont évalués sur une échelle en 10 points plutôt qu'en 4 points comme c'était le cas dans notre étude. En effet, une plus grande hétérogénéité de l'échantillon a tendance à influencer positivement la magnitude des corrélations obtenues<sup>[28]</sup>. Enfin, tel que mentionné en introduction, la moyenne d'âge des étudiants en médecine au Québec est inférieure à celle de ceux en contexte anglophone. En ce sens, les habiletés qualifiées de non-cognitives sont peut-être plus stables chez des étudiants plus âgés, ce qui expliquerait une plus grande validité prédictive des études en contexte anglophone.

La principale limite associée à cette étude est que nous avons évalué la validité prédictive des

MEMFI uniquement pour une des trois universités qui utilisent cet outil de sélection. Afin d'obtenir un portrait plus juste, il serait pertinent d'évaluer la validité prédictive des MEMFI pour les deux autres universités également. La limite suivante concerne la répartition des étudiants composant l'échantillon de l'externat. En effet, les candidats provenant de l'année d'admission 2009 sont surreprésentés ( $n = 201$ ) par rapport à ceux provenant des trois autres années d'admission, à savoir celles de 2010 ( $n = 41$ ), 2011 ( $n = 0$ ) et 2012 ( $n = 0$ ). La cohorte 2009 représente ainsi 83 % de l'échantillon global de l'externat ( $n = 242$ ). L'année 2009 est la première année d'utilisation des MEMFI par ces trois universités et certains ajustements ont été apportés depuis, tels que l'augmentation du nombre de stations. La validité prédictive des MEMFI pourrait donc être différente, possiblement supérieure, pour les MEMFI des années subséquentes. Par ailleurs, les résultats aux ECOS du programme de médecine étaient confondus avec ceux aux cours intégrateurs. Il était alors impossible d'estimer la force de la corrélation entre les MEMFI et les ECOS afin de les comparer avec les corrélations rapportées par Eva et al.<sup>[19]</sup> ainsi que par Husbands et Dowell<sup>[21]</sup>. Enfin, nous avons étudié la validité prédictive des MEM en contexte francophone uniquement au Québec. Il faut donc être prudent avant de généraliser les résultats à d'autres contextes francophones.

## Conclusion

La présente étude avait pour objectif d'évaluer la validité prédictive des MEM en contexte francophone. Elle démontre que les MEMFI, qui sont les MEM développées conjointement par les trois facultés de médecine francophones québécoises et administrées dans quatre sites, sont un outil de sélection des candidats permettant de prédire les résultats aux études en médecine. Elle contribue à la littérature existante sur les MEM et les MEMFI en démontrant qu'elles font preuve de validité prédictive et qu'elles sont complémentaires à la performance académique

antérieure dans la sélection des candidats en médecine. Il appert donc valable d'utiliser les MEM en contexte francophone, et même multi-universités et multi-sites à l'instar des MEMFI, comme outil de sélection des candidats de manière à prédire en partie leur réussite aux études en médecine. L'évaluation de la validité prédictive des MEM en contexte francophone pourra être poussée plus loin en reprenant cette étude avec davantage de données pour les candidats admis après 2009. Il serait aussi souhaitable d'étudier la capacité prédictive des MEMFI sur les résultats aux examens de certification canadiens (EACMC parties I et II), à l'instar des études d'Eva et al.<sup>[8, 14]</sup>.

## Contributions

Jean-Sébastien Renaud a contribué à la conception du protocole de recherche, au recueil des données, aux analyses statistiques, à l'interprétation des résultats, ainsi qu'à l'écriture, la révision et l'approbation du manuscrit. Adrien Cantat a contribué aux analyses statistiques, à l'interprétation des résultats, ainsi qu'à l'écriture, la révision et l'approbation du manuscrit. Sawsen Lakhal a contribué à la conception du protocole de recherche, au recueil des données, aux analyses statistiques, à l'interprétation des résultats, ainsi qu'à la révision et l'approbation du manuscrit. Martine Bourget a contribué au recueil des données ainsi qu'à la révision et l'approbation du manuscrit. Christina St-Onge a contribué à l'interprétation des résultats ainsi qu'à la révision et l'approbation du manuscrit.

## Remerciements

Nous tenons à remercier particulièrement Dr Jean-François Montreuil et Dre Lucie Rochefort, vice-doyens à la Faculté de médecine de l'Université Laval, M. Robert Gagnon, psychométricien au Centre de pédagogie appliquée aux sciences de la santé de

l'Université de Montréal, ainsi que l'ensemble des personnes impliquées dans les mini entretiens multiples francophones intégrées.

## Approbation éthique

Selon l'énoncé de politique des trois conseils<sup>[29]</sup>, la présente étude est considérée comme une activité d'assurance qualité et ne nécessite pas l'approbation d'un comité d'éthique de la recherche.

## Déclaration d'intérêts

Aucun auteur ne déclare de conflit d'intérêts en lien avec le contenu de cet article.

## Références

1. Frank JR, Snell LS, Sherbino J, Boucher A. Référentiel de compétences CanMEDS 2015 pour les médecins. Collège royal des médecins et chirurgiens du Canada, 2015.
2. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing. Washington (DC): American Educational Research Association, 2014.
3. Salvatori P. Reliability and Validity of Admissions Tools Used to Select Students for the Health Professions. *Adv Health Sci Educ Theory Pract* 2001;6: 159-75.
4. Norman G. Editorial – The Morality of Medical School Admissions. *Adv Health Sci Educ Theory Pract* 2004;9:79-82.
5. Willingham WW, Young JW, Morris MM. Success in college: The role of personal qualities and academic ability. New York: College Entrance Examination Board, 1985.
6. Eva KW, Rosenfeld J, Reiter HI, Norman GR. An admissions OSCE: the multiple mini-interview. *Med Educ* 2004;38:314-26.
7. Eva KW, Reiter HI, Rosenfeld J, Norman GR. The relationship between interviewers' characteristics and ratings assigned during a multiple mini-interview. *Acad Med* 2004;79:602-9.
8. Eva KW, Reiter HI, Rosenfeld J, Trinh K, Wood TJ, Norman GR. Association between a medical school admission process using the multiple mini-interview and national licensing examination scores. *JAMA* 2012;308:2233-40.
9. Harden RM, Gleeson FA. Assessment of clinical competence using an objective structured clinical examination (OSCE). *Med Educ* 1979;13:39-54.
10. Knorr M, Hissbach J. Multiple mini-interviews: same concept, different approaches. *Med Educ* 2014;48: 1157-75.
11. Roberts C, Walton M, Rothnie I, Crossley J, Lyon P, Kumar K, et al. Factors affecting the utility of the multiple mini-interview in selecting candidates for graduate-entry medical school. *Med Educ* 2008;42: 396-404.
12. Dore KL, Kreuger S, Ladhani M, Rolfson D, Kurtz D, Kulasegaram K, et al. The reliability and acceptability of the multiple mini-interview as a selection instrument for postgraduate admissions. *Acad Med* 2010;85:S60-S3.
13. Serres ML, Peeters MJ. Overcoming content specificity in admission interviews: The next generation? *Am J Pharm Educ* 2012;76:207.
14. Eva KW, Reiter HI, Trinh K, Wasi P, Rosenfeld J, Norman GR. Predictive validity of the multiple mini-interview for selecting medical trainees. *Med Educ* 2009;43:767-75.
15. Hecker K, Violato C. A generalizability analysis of a veterinary school multiple mini interview: Effect of number of interviewers, type of interviewers, and number of stations. *Teach Learn Med* 2011;23: 331-6.
16. Tavares W, Mausz J. Assessment of non-clinical attributes in paramedicine using multiple mini-interviews. *Emerg Med J* 2013;32:170-5.
17. Dowell J, Lynch B, Till H, Kumwenda B, Husbands A. The multiple mini-interview in the UK context: 3 years of experience at Dundee. *Med Educ* 2012;34: 297-304.
18. Griffin B, Harding DW, Wilson IG, Yeomans ND. Does practice make perfect? The effect of coaching



- and retesting on selection tests used for admission to an Australian medical school. *Med Educ* 2008;189:270-3.
19. Eva KW, Reiter HI, Rosenfeld J, Norman GR. The ability of the multiple mini-interview to predict pre-clerkship performance in medical school. *Acad Med* 2004;79:S40-S2.
  20. Reiter HI, Eva KW, Rosenfeld J, Norman GR. Multiple mini-interviews predict clerkship and licensing examination performance. *Med Educ* 2007;41:378-84.
  21. Husbands A, Dowell J. Predictive validity of the Dundee multiple mini-interview. *Med Educ* 2013;47:717-25.
  22. Bergeron L, St-Onge C, Martel S, Hanna D. Évaluation éducatrice d'un dispositif d'entretiens structurés multiples pour la sélection de candidats dans un programme postgradué de dermatologie. *Pédagogie Médicale* 2011;12:17-27.
  23. St-Onge C, Côté DJ, Brailovsky C. Utilisation des Mini Entretiens Multiples en contexte francophone : étude de généralisabilité. *Mesure et Évaluation en Éducation* 2009;32:49-69.
  24. Gagnon R. Les entretiens structurés multiples pour la sélection des candidats aux programmes de formation en médecine. *Pédagogie Médicale* 2011;12:5-6.
  25. van der Vleuten CPM, Verwijnen GM, Wijnen WHFW. Fifteen years of experience with progress testing in a problem-based learning curriculum. *Med Teach* 1996;18:103-9.
  26. Scallion G. L'évaluation des apprentissages dans une approche par compétences. Saint-Laurent: Éditions du Renouveau pédagogique, 2004.
  27. Carifio J, Perla R. Resolving the 50-year debate around using and misusing Likert scales. *Med Educ* 2008;42:1150-2.
  28. Steinberg WJ. *Statistics alive!* Thousand Oaks: Sage Publications, 2008.
  29. Conseil de recherche en sciences humaines du Canada, Conseil de recherche en sciences naturelles et en génie du Canada, Instituts de recherche en santé du Canada. Énoncé de politique des trois Conseils : Éthique de la recherche avec des être humains. Ottawa, Canada: Secrétariat interagences en éthique de la recherche, décembre 2010.
- 
- Correspondance et offprints :** Jean-Sébastien Renaud. Faculté de médecine, Pavillon Ferdinand-Vandry, local 2881-D Université Laval, 1050 avenue de la Médecine. Québec (Québec) G1V 0A6. Mailto : jean-sebastien.renaud@fmed.ulaval.ca