

# Association Between a Medical School Admission Process Using the Multiple Mini-interview and National Licensing Examination Scores

Kevin W. Eva, PhD

Harold I. Reiter, MD

Jack Rosenfeld, PhD

Kien Trinh, MD

Timothy J. Wood, PhD

Geoffrey R. Norman, PhD

**M**ODERN CONCEPTIONS OF medical practice demand more of practitioners than a strong knowledge base.<sup>1</sup> By emphasizing compassionate care, professionalism, and interpersonal skill, the Accreditation Council for Graduate Medical Education core competencies indicate that physicians are expected to possess strong personal qualities distinct from academic achievement.<sup>2</sup> There is evidence of a link between these aspects of practice and quality of care.<sup>3-7</sup> At the level of medical school admissions, incoming grade point average (GPA) and Medical College Admission Test results have been found to be reasonably good determinants of academic success.<sup>8-10</sup> More problematic, however, has been the identification of measures capable of predicting nonacademic success despite the considerable resources most medical schools allocate to interviewing applicants.<sup>9</sup>

Most validity studies in this domain are correlational in nature. Interpretation of the resulting statistics is difficult because restriction of range low-

**Context** There has been difficulty designing medical school admissions processes that provide valid measurement of candidates' nonacademic qualities.

**Objective** To determine whether students deemed acceptable through a revised admissions protocol using a 12-station multiple mini-interview (MMI) outperform others on the 2 parts of the Canadian national licensing examinations (Medical Council of Canada Qualifying Examination [MCCQE]). The MMI process requires candidates to rotate through brief sequential interviews with structured tasks and independent assessment within each interview.

**Design, Setting, and Participants** Cohort study comparing potential medical students who were interviewed at McMaster University using an MMI in 2004 or 2005 and accepted (whether or not they matriculated at McMaster) with those who were interviewed and rejected but gained entry elsewhere. The computer-based MCCQE part I (aimed at assessing medical knowledge and clinical decision making) can be taken on graduation from medical school; MCCQE part II (involving simulated patient interactions testing various aspects of practice) is based on the objective structured clinical examination and typically completed 16 months into postgraduate training. Interviews were granted to 1071 candidates, and those who gained entry could feasibly complete both parts of their licensure examination between May 2007 and March 2011. Scores could be matched on the examinations for 751 (part I) and 623 (part II) interviewees.

**Intervention** Admissions decisions were made by combining z score transformations of scores assigned to autobiographical essays, grade point average, and MMI performance. Academic and nonacademic measures contributed equally to the final ranking.

**Main Outcome Measures** Scores on MCCQE part I (standardized cut-score, 390 [SD, 100]) and part II (standardized mean, 500 [SD, 100]).

**Results** Candidates accepted by the admissions process had higher scores than those who were rejected for part I (mean total score, 531 [95% CI, 524-537] vs 515 [95% CI, 507-522];  $P=.003$ ) and for part II (mean total score, 563 [95% CI, 556-570] vs 544 [95% CI, 534-554];  $P=.007$ ). Among the accepted group, those who matriculated at McMaster did not outperform those who matriculated elsewhere for part I (mean total score, 524 [95% CI, 515-533] vs 546 [95% CI, 535-557];  $P=.004$ ) and for part II (mean total score, 557 [95% CI, 548-566] vs 582 [95% CI, 569-594];  $P=.003$ ).

**Conclusion** Compared with students who were rejected by an admission process that used MMI assessment, students who were accepted scored higher on Canadian national licensing examinations.

*JAMA.* 2012;308(21):2233-2240

www.jama.com

**Author Affiliations:** Department of Medicine, University of British Columbia, Vancouver, Canada (Dr Eva); Departments of Oncology (Dr Reiter), Pathology and Molecular Medicine (Dr Rosenfeld), Family Medicine (Dr Trinh), and Clinical Epidemiology and Biostatistics (Dr Norman), McMaster University, Hamilton,

Ontario, Canada; and Department of Medicine, University of Ottawa, Ottawa, Ontario (Dr Wood).

**Corresponding Author:** Kevin W. Eva, PhD, Centre for Health Education Scholarship, JPPN 3324, 910 W 10th Ave, Vancouver, BC V5Z 4E3, Canada (kevin.eva@ubc.ca).

**For editorial comment see p 2250.**

ers the correlations that can be observed within a pool of accepted applicants and because it is impossible to determine how rejected applicants would have performed had they been admitted.<sup>10</sup> A cohort study by Kulatunga-Moruzi and Norman<sup>10</sup> found that applicants who were accepted through an extensive screening process that included autobiographical essays, personal interviews, and tutorial simulations were indistinguishable on graduation (according to national licensing examination scores) from those who were rejected but gained entry elsewhere. We studied whether or not an admission process that instead uses the multiple mini-interview (MMI),<sup>11</sup> a series of independent observations collected via sequential structured interviews, would yield good prediction of national licensing examination scores.

## METHODS

Approval of this study was received from the Faculty of Health Sciences research ethics board at McMaster University and Medical Council of Canada (MCC). Because all data were preexisting and deidentified after a third party merged data sets, explicit informed consent from individual participants was not required.

### Study Population

Participants were those who interviewed for the McMaster undergraduate doctor of medicine (MD) program in the 2004 and 2005 admissions cycles, the first 2 cohorts to be admitted using the revised protocol. Because McMaster offers a 3-year training program, students entered the classes of 2007 and 2008. All but 1 of the other 16 medical schools in Canada offer 4-year training programs. As a result, applicants from the same admissions cycles who attended a different school generally graduated in 2008 or 2009.

Graduates of Canadian medical schools are eligible to take the MCC Qualifying Examination (MCCQE) part I in May of their graduation year. They must then complete a year of

clinical postgraduate training to be eligible for part II of the examination. As a result, Canadian graduates typically complete their licensing examination 16 to 20 months after completing their MD, and those in the cohorts considered for this study generally completed both parts of the MCCQE by March 2011. Deidentified data from McMaster University and MCC were merged in the summer of 2011 by a third party following MCC guidelines for sharing examination data, a process that ensures anonymity of the participants and their scores. Sex and age of participants were available to facilitate the merging of data, but age was deleted postmerge to ensure anonymity of the data set. No other personal information was available on participants or their demographics. Only those who had completed at least part I of the MCCQE by June 2011 were included in the study.

### Admissions Protocols

As of 2004, the MD program at McMaster abandoned its traditional use of both a panel-style personal interview, in which 3 examiners interviewed each candidate for a period of 25 minutes (followed by 30 minutes of scoring), and a simulated tutorial in which candidates were placed in groups of 6 or 7 and observed interacting as they worked through a presented problem. Dominant concerns leading to this decision were poor test-retest reliability and weak predictive validity.<sup>12</sup> Instead, applicants were invited to interview based on undergraduate GPA weighted approximately 70% and an autobiographical submission weighted approximately 30%, a protocol that was unchanged from preceding years. Those invited to interview were then assessed using a 12-station MMI.<sup>11</sup>

The MMI consists of a series of brief interviews modeled after the objective structured clinical examination (OSCE), in which candidates typically have 8 minutes to discuss an issue with an interviewer, demonstrate the capacity to work through a challenging interpersonal situation

presented by an actor, or perform a task with another candidate. Previously reported examples include the need to counsel a colleague who was afraid of flying after the September 11 attacks and a discussion of the ethical principles involved in deciding whether a physician should treat a patient with a known placebo for the sake of offering reassurance.<sup>11</sup> Examiners are then given 2 minutes to score the applicant's performance while the applicant prepares for the next station.

McMaster uses a 12-station MMI that is designed to dominantly focus on ethical issues, communication, and collaborative tasks.<sup>13</sup> Because few stations are used within any subdomain, subdomain scores cannot be considered reliable metrics. A variety of studies have demonstrated that aggregating across 12 stations routinely offers interstation reliability in the 0.7 to 0.8 range.<sup>12</sup> This is substantially higher than the inter-interview reliability of less than 0.4 observed in panel-based interviews.<sup>14</sup> The decision to change to this process was based on this evidence of improved reliability and lower resource requirements relative to the traditional interview process.<sup>15</sup> Since then, the MMI has been shown to improve the association between admissions data and clinical performance during clerkship and the ethical reasoning/communication components of the national licensing examinations.<sup>12,16</sup>

In 2004, McMaster was the only Canadian medical school to use an MMI, and in 2005 a second medical school from a different province (hence, sharing few applicants with McMaster) began using the MMI, thus enabling us to make an uncontaminated comparison between students accepted via an MMI and those rejected but accepted elsewhere using traditional interview procedures.

After applicants completed the MMI, scores on all measures underwent *z* score transformation and were combined so that approximately 30% of the

weight was placed on GPA and 70% on the MMI. The final admission decision was made by rank order with minimal file review conducted by the admissions committee to identify exceptional circumstances, such as egregious unprofessional behaviors.

### Outcome Variables

To gain a license to practice medicine in Canada, a physician must hold a valid medical degree and meet the requirements of the provincial regulatory authorities, which usually includes being a licentiate of MCC (granted to those who pass both parts of the MCCQE). The MCCQE part I is a computer-based examination that includes an adaptive multiple-choice question component testing general medical knowledge and a short-answer key-features style assessment of clinical decision-making skills. This examination is generally taken in May immediately following completion of training for an MD. Scoring of the multiple-choice question component is conducted using item response theory, with both components combined and transformed to a scale with a fixed cut-score of 390 and SD of 100 so that performance is comparable from one year to the next.<sup>17</sup> Scores from this examination have been found to be predictive of a variety of clinical practice quality indicators.<sup>5,6</sup> For the purpose of conveying feedback to candidates (not for decision making), the MCC generates subscores for specific clinical rotations and for considerations of legal, ethical, and organizational (CLEO) aspects of practice.

The MCCQE part II is an OSCE consisting of 14 stations, each involving interaction with a standardized patient.<sup>18</sup> This examination is generally completed 16 months into residency training. Physician examiners observe the interaction and score candidates using a combination of station-specific checklists and rating scales. A modified borderline groups method is used to determine the cut-score on each station,<sup>19</sup> and scores are transformed to the same scale each year (mean=500, SD=100). Scores from this examina-

tion have been found to be related to patient complaints and clinical practice quality indicators.<sup>4</sup> For feedback purposes only, subscores are generated. On part II, these subscores are data gathering, problem solving, patient interaction (which includes communication skills), and CLEO.

### Statistical Analysis

The unit of analysis was the examinee. Standard descriptive statistics were used to summarize the data. The primary analyses were univariate analyses of variance performed on MCCQE scores to look for differences between candidates accepted to McMaster (whether they matriculated at McMaster or elsewhere) and those who were rejected but gained entry to another medical school. These analyses were then repeated by dividing the "accepted at McMaster" group into those who attended McMaster and those who completed their medical training elsewhere. Multiple regression analyses were also performed on each dependent variable, including GPA, autobiographical submission, and MMI scores as predictor variables to determine the extent to which admission scores were independently related to the outcomes of interest. Because the sample size was determined by the interview process, an a priori power analysis was not conducted. All tests were 2-sided with  $P < .05$  used as the significance threshold. All analyses were performed using PASW Statistics 18.0 for Macintosh (IBM SPSS).

### RESULTS

In 2004 and 2005, 6049 applicants applied for medical school admission to McMaster. Of 1071 applicants who were brought to interview, 521 (48.6%) were offered admission via the revised admissions process. TABLE 1 presents the demographic characteristics of all interviewed applicants. Six hundred fifty-three of 1071 applicants (61.0%) were women, a proportion that was comparable in both accepted and rejected interviewees. The mean age was 24.0 years, with accepted applicants

being slightly older than rejected interviewees (mean, 24.3 vs 23.8 years, respectively).

A total of 70.1% (751/1071) of all interviewees were found in the MCC database and had their admissions scores matched to MCCQE part I scores. Part II scores were available for 623 of these 751, the number likely lower because not all had completed part II by the time the data were collected for this study. Table 1 illustrates that the matched sample included 90.6% (472/521) of all accepted candidates and 50.7% (279/550) of all rejected candidates. Given the successful match rate for the former group, it seems likely that almost all nonmatched candidates from the latter group did not gain entry to a Canadian medical school in the years considered in this study. Of the 472 accepted candidates who matched on part I scores, 274 (58.1%) matriculated at McMaster, 128 (27.1%) matriculated elsewhere, and 70 (14.8%) could not be matched to site of matriculation. Sex did not differ as a function of acceptance in the matched sample, and age could not be considered because it was excluded from the deidentified data to reduce the potential that individual candidates could be recognized.

### Admissions Results

TABLE 2 presents the scores achieved on each of the admissions instruments (GPA, autobiographical submission, and MMI) as a function of whether or not applicants were accepted to McMaster. A statistically significant difference was observed between accepted and rejected interviewees for GPA and MMI scores, expected given that the instruments defined who was accepted. Although the association with autobiographical submission had a moderate effect size, this comparison was not statistically significant, likely because lesser weight was applied to this variable in the admissions decision. The correlation between GPA and MMI was  $r=0.006$ ; between GPA and autobiographical submission,  $r=-0.38$ ; and

between MMI and autobiographical submission,  $r=0.014$ .

**Relation Between Admissions Decisions and Licensing Examination Performance**

Those candidates who were accepted by McMaster achieved mean performance scores on the MCCQE part I that were greater than the scores achieved by rejected candidates (TABLE 3). That was true for the total mean score (531 [95% CI, 524-537] vs 515 [95% CI, 507-522], respectively;  $P=.003$ ) and the

multiple-choice question (mean score, 542 [95% CI, 535-549] vs 529 [95% CI, 521-536],  $P=.02$ ) and clinical decision making (mean score, 467 [95% CI, 461-474] vs 444 [95% CI, 435-454],  $P<.001$ ) components of the examination. There was no statistically significant association between the CLEO subscore and status as an accepted vs rejected applicant (mean score, 526 [95% CI, 519-533] vs 520 [95% CI, 511-529],  $P=.17$ ).

On the MCCQE part II, there were statistically significant differences fa-

voring participants accepted to McMaster for the total mean score (563 [95% CI, 556-570] vs 544 [95% CI, 534-554], respectively;  $P=.007$ ), the mean CLEO subscore (553 [95% CI, 546-560] vs 520 [95% CI, 509-531];  $P<.001$ ), and the mean patient interaction subscore (which includes communication skills; 560 [95% CI, 554-567] vs 538 [95% CI, 526-549];  $P=.002$ ) (TABLE 4). The data gathering and problem solving subscores did not differ across the 2 groups (mean scores, 541 [95% CI, 534-549] vs 541 [95% CI, 528-

**Table 1.** Characteristics of Applicants Interviewed for Admission to the Undergraduate MD Program at McMaster University in 2004 or 2005 (N = 1071) and Subset Who Could Be Matched to Performance on the MCCQE (n = 751)

Variable	Interviewees			Statistics <sup>a</sup>
	All (N = 1071)	Accepted (n = 521)	Rejected (n = 550)	
Year of interview, No.				
2004	376	234	142	
2005	695	287	408	
Sex, No. (%)				$\chi^2 = 1.68, P = .19$
Men	418 (39.0)	193 (37.0)	225 (40.9)	
Women	653 (61.0)	328 (63.0)	325 (59.1)	
Age, mean (SD) [95% CI], y	24.0 (3.5) [23.8-24.2]	24.3 (3.9) [23.9-24.6]	23.8 (3.04) [23.6-24.1]	$t = 2.0, P = .046$
Interviewees matched to MCCQE scores, No./total No. (%)	751/1071 (70.1) <sup>b</sup>	472/521 (90.6)	279/550 (50.7)	
Year of graduation, No. <sup>c</sup>				
2007	135	133	2 <sup>d</sup>	
2008	247	207	40	
2009	287	126	161	
2010	82	6	76	
Sex, No. (%)				$\chi^2 = 2.45, P = .12$
Men	301 (40.1)	179 (37.9)	122 (43.7)	
Women	450 (59.9)	293 (62.1)	157 (56.3)	

Abbreviations: MCCQE, Medical Council of Canada Qualifying Examination; MD, doctor of medicine.  
<sup>a</sup> $P$  values reflect comparison between interviewees who were accepted by the McMaster undergraduate MD admissions protocol and those who were rejected.  
<sup>b</sup>Ten of the matched participants graduated from non-Canadian medical schools (1.3% of the total sample).  
<sup>c</sup>McMaster is 1 of 2 medical schools in Canada that graduates students after 3 years. All other schools offer 4-year programs.  
<sup>d</sup>This number is small because only 1 other school in Canada has a 3-year undergraduate MD (ie, few rejected applicants can be expected to have written part I of the licensing examination within 3 years).

**Table 2.** Comparison of McMaster-Based Admissions Data for Participants Whose Admissions Data Could Be Matched to Performance on the MCCQE

Variable	Mean (SD) [95% CI]			Statistics <sup>a</sup>
	All (n = 751)	Accepted (n = 472)	Rejected (n = 279)	
Grade point average <sup>b</sup>	3.82 (0.14) [3.81-3.83]	3.85 (0.13) [3.83-3.86]	3.78 (0.14) [3.76-3.79]	$t = 5.93, P < .001; d = .62$
Autobiographical submission <sup>c</sup>	15.0 (2.43) [14.8-15.2]	15.1 (2.38) [14.9-15.4]	14.8 (2.49) [14.5-15.1]	$t = 1.33, P = .18; d = .44$
Multiple mini-interview scores <sup>d</sup>	66.4 (12.20) [65.5-67.2]	70.5 (10.87) [69.6-71.5]	59.4 (11.06) [58.1-60.6]	$t = 11.08, P < .001; d = .52$

Abbreviations: MCCQE, Medical Council of Canada Qualifying Examination; MD, doctor of medicine.  
<sup>a</sup> $P$  values and  $d$  values reflect comparison between interviewees who were accepted by the McMaster undergraduate MD admissions protocol (whether or not they matriculated at McMaster) and those who were rejected.  $d$  refers to Cohen  $d$ , the difference between means divided by the pooled standard deviation.  
<sup>b</sup>Grade point average was calculated by averaging all undergraduate coursework when converted to a 4-point scale.  
<sup>c</sup>Autobiographical submission scores represent the sum total of 3 raters' opinions regarding how well candidates responded to 5 questions submitted with candidates' application package and scored using 7-point scales (minimum=3, maximum=21).  
<sup>d</sup>Multiple mini-interview scores represent the sum total of 12 raters' opinions regarding how well candidates performed across 12 stations, scored using a 10-point scale (minimum=12, maximum=120).

553],  $P = .84$  and 538 [95% CI, 531-546] vs 527 [95% CI, 515-539],  $P = .14$ , respectively).

To confirm that the differences reported did not arise from McMaster's curriculum outperforming those of the other schools, we examined each of the dependent measures as a function of whether accepted applicants enrolled at McMaster or enrolled elsewhere. Those accepted candidates who enrolled at McMaster did not outperform those who enrolled elsewhere on any outcome. For part I, total mean scores were 524 (95% CI, 515-533) and 546 (95% CI, 535-557;  $P = .004$ ), and for part II, total mean scores were 557 (95% CI, 548-566) and 582 (95% CI, 569-594;  $P = .003$ ). Subscores for part I and part II are illustrated in eTable 1 and eTable 2 (available at <http://www.jama.com>).

A series of regression analyses was performed to determine the extent to which individual admissions measures were related to the various outcomes of interest (TABLE 5). In general, autobiographical submission scores were unrelated to performance

on the MCCQE examinations. Grade point average was particularly predictive of MCCQE part I scores. In contrast, MMI was independently predictive of all of the dependent measures considered. Thus, the MMI showed incremental validity in combination with the other variables available at admission for both MCCQE part I and part II.

**COMMENT**

To our knowledge, this study is the first to assess a full graduating class that was selected using the MMI. This study is now difficult to replicate in Canadian medical schools because 14 of the 17 programs currently use the MMI or an adaptation of the same principles. In the United States, systematic data are not available, but use of the MMI has been reported in dentistry<sup>20</sup> and physician assistant programs.<sup>21</sup> According to records kept by the Association for American Medical Colleges and ProFitHR, it appears that at least 22 US medical schools are known to be using the MMI or an MMI hybrid as

of October 2012 (Kirch D, President, Association of American Medical Colleges, written communication, October 18, 2012. Snelgrove T, President, ProFitHR, written communication, October 17, 2012). Combined with the Canadian schools, that means nearly one-quarter of schools (22.8%, 36/158) accredited by the Liaison Committee on Medical Education are known to use an MMI as part of their admissions process.

We found reasonably stable differences on 2 parts of a national licensing examination indicating that those accepted by the MMI-based admissions process outperformed those who were rejected but gained entry elsewhere. These differences arose 4 to 6 years after the selection decision was made and cannot be attributed solely to the curricular success of the university, because those who were accepted but chose to study elsewhere performed as well as those who matriculated at McMaster. It is important to note that McMaster's curriculum has changed since the time of this study and that, as a result, any differences ob-

**Table 3.** Comparison of MCCQE Part I Data for Participants Whose McMaster-Based Admissions Data Could Be Matched to Part I Performance

MCCQE Part I Scores	Mean (SD) [95% CI]			Statistics <sup>a</sup>
	All (n = 751)	Accepted (n = 472)	Rejected (n = 279)	
Total	525 (70.4) [520-530]	531 (72.1) [524-537]	515 (66.3) [507-522]	$F = 8.3, P = .003; d = .24$
Multiple choice	537 (72.9) [532-542]	542 (76.8) [535-549]	529 (65.0) [521-536]	$F = 5.40, P = .02; d = .19$
Clinical decision making	459 (78.3) [453-464]	467 (74.4) [461-474]	444 (82.5) [435-454]	$F = 13.2, P < .001; d = .29$
CLEO subscore	524 (79.3) [518-529]	526 (81.3) [519-533]	520 (75.8) [511-528]	$F = 1.9, P = .17; d = .08$

Abbreviations: CLEO, considerations of ethical, legal, and organizational aspects of practice; MCCQE, Medical Council of Canada Qualifying Examination; MD, doctor of medicine. <sup>a</sup> $P$  values and  $d$  values reflect comparison between interviewees who were accepted by the McMaster undergraduate MD admissions protocol (whether or not they matriculated at McMaster) and those who were rejected.  $d$  refers to Cohen  $d$ , the difference between means divided by the pooled standard deviation.

**Table 4.** Comparison of MCCQE Part II Data for Participants Whose McMaster-Based Admissions Data Could Be Matched to Part II Performance

MCCQE Part II Scores	Mean (SD) [95% CI]			Statistics <sup>a</sup>
	All (n = 623)	Accepted (n = 437)	Rejected (n = 186)	
Total	557 (73.3) [552-563]	563 (73.0) [556-570]	544 (72.5) [533-554]	$F = 7.2, P = .007; d = .26$
Data gathering subscore	541 (84.5) [534-548]	541 (83.1) [534-549]	541 (87.9) [528-553]	$F = 0.04, P = .84; d = .01$
Problem solving subscore	535 (83.0) [528-541]	538 (82.3) [531-546]	527 (84.2) [515-539]	$F = 2.2, P = .14; d = .14$
Patient interaction subscore	553 (73.2) [548-559]	560 (69.2) [554-567]	538 (79.8) [526-549]	$F = 10.1, P = .002; d = .30$
CLEO subscore	543 (79.1) [537-549]	553 (77.6) [546-560]	520 (77.9) [509-531]	$F = 20.2, P < .001; d = .42$

Abbreviations: CLEO, considerations of ethical, legal, and organizational aspects of practice; MCCQE, Medical Council of Canada Qualifying Examination; MD, doctor of medicine. <sup>a</sup> $P$  values and  $d$  values reflect comparison between interviewees who were accepted by the McMaster undergraduate MD admissions protocol (whether or not they matriculated at McMaster) and those who were rejected.  $d$  refers to Cohen  $d$ , the difference between means divided by the pooled standard deviation.

**Table 5.** Results of Linear Regression Analyses of the Relationship Between MCCQE Scores (Part I and Part II) and Grade Point Average, Autobiographical Submission, and Multiple Mini-Interview

Dependent Variable	Predictor Variables, Standardized $\beta$ (P Values) <sup>a</sup>			Multiple R <sup>b</sup>
	GPA	ABS	MMI	
MCCQE part I scores				
Total score	0.22 (P < .001)	-0.08 (P = .03)	0.12 (P < .001)	0.29
Multiple choice	0.23 (P < .001)	-0.08 (P = .03)	0.10 (P < .001)	0.29
Clinical decision making	0.11 (P = .01)	-0.05 (P = .17)	0.13 (P < .001)	0.19
CLEO	0.10 (P = .01)	0.10 (P = .81)	0.10 (P = .01)	0.14
MCCQE part II scores				
Total	0.09 (P = .03)	-0.06 (P = .18)	0.21 (P < .001)	0.25
Data gathering	0.08 (P = .06)	-0.01 (P = .88)	0.10 (P = .01)	0.13
Problem solving	0.08 (P = .08)	-0.04 (P = .33)	0.11 (P = .01)	0.15
Patient interaction	0.00 (P = .99)	-0.08 (P = .06)	0.23 (P < .001)	0.24
CLEO	-0.01 (P = .81)	-0.05 (P = .22)	0.22 (P < .001)	0.23

Abbreviations: ABS, autobiographical submission; CLEO, considerations of legal, ethical, and organizational aspects of practice; GPA, grade point average; MCCQE, Medical Council of Canada Qualifying Examination; MMI, multiple mini-interview.

<sup>a</sup>Standardized  $\beta$  weights are estimates of the relationship between 2 variables that have been standardized to have a variance equal to 1. As such, they describe how many standard deviations a dependent variable will change per standard deviation of the predictor variable.

<sup>b</sup>Multiple R provides the correlation between the model (taking all 3 predictor variables into account) and the dependent measures.

served between curricula are no longer reflective of current practice.<sup>22</sup>

The significant prediction of licensing examination performance observed in this study contrasts with an earlier study<sup>10</sup> reporting that students accepted using the previous McMaster admissions process, which involved panel-based interviews and simulated tutorials, were indistinguishable from those who were rejected when compared against the same outcome measures. The MMI appears to show a stronger relationship with measures of clinical skills and ethical reasoning when directly compared against traditional panel-style interview protocols.<sup>12,16,23</sup> The inverse relationship between GPA and autobiographical submission was unanticipated but fits with previous work indicating that the relationship between academic and nonacademic measures of performance moves from negative to positive as the measures become increasingly focused on specific domains of performance (ie, as they move from generic and broad measures to measures of medical competencies).<sup>12</sup>

Gaining entry to an MD training program is the single greatest determinant of who will practice medicine in North America as more people are excluded during the admissions process than at any other stage in training. Al-

though accepted applicants generally go on to succeed at passing the licensing examination, this is only an indicator of minimal competence rather than success of the admissions protocol. Moreover, during the admissions process, high-stakes decisions are being made on behalf of applicants, so the selection process should be as fair and meaningful as possible.<sup>24</sup>

No selection process will be perfect. Many factors preclude any admissions protocol from predicting relevant outcomes with complete accuracy, including gamesmanship, the curriculum, natural maturation, and the complex and probabilistic nature of life itself. However, the empirical evidence supports considerable room for improvement. In European countries in which students have traditionally been admitted based on self-selection (either via open admissions practices in which students are allowed to choose their domain of study without institutional selection or through lottery-based selection), improvement in performance and graduation rate has been observed after academic measures were implemented to create a barrier that had to be overcome (typically using GPA or standardized tests for the sake of selection decision making).<sup>25-27</sup>

Personal interviews have become the standard through which health profes-

sional schools assess nonacademic aspects of applicants.<sup>28</sup> There is a broad spectrum of qualities that can be used to select candidates,<sup>29</sup> but traditional interviews may not be able to identify the top candidates.<sup>10</sup> At issue is not the skill or intent of the interviewers, but rather the consistency of performance of the candidates. A single interview will generate an impression of interpersonal skill, thoughtfulness, and general demeanor in that particular interview. However, in a different interview setting, the performance may change significantly: the effect of occasion is generally a greater source of error in interview measurement than is the effect of differences of opinion between raters.<sup>11,14</sup>

There are 3 main limitations of this study. First, the outcomes examined are not direct measures of actual practice. Subscores on the MCCQE parts I and II need to be interpreted with caution because they are based on only a subset of items and hence have lower reliability. With that caveat, the communication skills subscore and the part I and part II total scores are arguably important measures; performance on these examinations has been linked to real-world practice differences deemed indicative of the quality of health care provided by Canadian physicians.<sup>4-6</sup> As the first cohorts admitted to any profes-

sional school via the MMI are just now reaching the stage of independent practice, more time must pass before directly testing the link between admissions decisions and clinical care. Second, the sample size and lack of additional data on which to compare various applicant demographics precludes conducting additional analyses that would be informative.

Third, the comparison used to claim that replacement of traditional panel-style interviews with an MMI yielded improvement in predictive ability is a historical one. Although the acceptance decision using an MMI-based admissions process provided better discrimination than was seen in the same institution previously,<sup>10</sup> the lack of a current control condition prevents absolute comparative conclusions from being drawn. The findings should not be interpreted as indicating that the MMI, as implemented at McMaster, is necessarily the best admissions approach or even the only approach that would work well. Rather, the data presented here should most conservatively be treated as proof of concept of the value of performing quality assurance analyses and striving for continuous quality improvement within the context of any institution's admissions protocol. Whether the logistical requirements necessary for mounting an MMI add or remove burden from an institution's resource use will depend on the institution's current practices.<sup>15</sup>

The MMI should be considered a process of assessment, not a tool or an instrument. Just as a multiple-choice examination can be populated with questions representative of diverse content areas, an MMI can be populated with highly variable stations. What is measured and what outcomes can be predicted will depend heavily on what qualities an institution designs its MMI to address and what weight is assigned to each measure collected. The McMaster MMI was designed to emphasize ethical and interpersonal dilemmas.<sup>13</sup> Others have chosen to emphasize different things and therefore

may need to look at different outcomes to determine their success.

Regardless of an institution's focus, the key issue that both defines the MMI and strengthens its psychometric properties is the repeated sampling of performance. Feasibly sampling candidate performance over multiple situations requires relatively brief encounters. There may not be an advantage to longer interviews, because examiners tend to form impressions very quickly, with additional time being largely redundant.<sup>30,31</sup> In addition, the more time a candidate spends with an interviewer, the greater the opportunity to divert the conversation to issues that are distinct from the intended focus of the interview, which tends to add error variance in a manner that can harm rather than help the measurements collected.<sup>32</sup>

In conclusion, there appears to be a complementary relationship between GPA and the MMI process, with the former related to more knowledge-oriented outcome measures and the latter to more clinical/ethical/interpersonal skill-oriented outcomes. Our study demonstrates that at McMaster, a GPA plus MMI approach has yielded better outcomes than were achieved by the historical use of GPA plus panel-style interview/simulated tutorial. Future research should examine characteristics of individual MMI stations to determine if some implementation strategies are more effective than others. Furthermore, research into how various weighting schemes influence the composition of the student population may be beneficial; however, the appropriate choice of weighting scheme may be context specific as long as reliable and trustworthy measures are being weighted.

**Author Contributions:** Dr Eva had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

*Study concept and design:* Eva, Reiter, Norman.

*Acquisition of data:* Eva, Rosenfeld, Trinh, Wood.

*Analysis and interpretation of data:* Eva.

*Drafting of the manuscript:* Eva.

*Critical revision of the manuscript for important intellectual content:* Eva, Reiter, Rosenfeld, Trinh, Wood, Norman.

*Statistical analysis:* Eva.

*Administrative, technical, or material support:* Eva, Reiter, Rosenfeld, Trinh, Wood, Norman.

**Conflict of Interest Disclosures:** All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Drs Eva, Trinh, and Norman reported having no conflicts of interest beyond testing the effectiveness of an educational innovation that they have been involved in creating. Drs Reiter and Rosenfeld reported being stakeholders in ProFitHR, a commercial enterprise that markets their experience to assist groups in mounting multiple mini-interviews. They also receive royalties through McMaster University's licensing of MMI questions. Dr Wood reported being an MCC employee at the time this study was conducted.

**Disclaimer:** This article does not necessarily reflect MCC policy, and MCC provides no official endorsement.

**Online-Only Material:** eTables 1 and 2 are available at <http://www.jama.com>.

**Additional Contributions:** We are grateful for the support of Steve Slade, BA, CAPER (Canadian Post-M.D. Education Registry), for his assistance in merging the data from both institutions involved in this study and Marguerite Roy, PhD, of the MCC, for her assistance with the construction of data files. Neither received compensation for their contributions.

## REFERENCES

1. Eva KW, Lohfeld L, Dhaliwal G, Mylopoulos M, Cook DA, Norman GR. Modern conceptions of elite medical practice among internal medicine faculty members. *Acad Med.* 2011;86(10)(suppl):S50-S54.
2. Swing SR, Clyman SG, Holmboe ES, Williams RG. Advancing resident assessment in graduate medical education. *J Grad Med Educ.* 2009;1(2):278-286.
3. Papadakis MA, Teherani A, Banach MA, et al. Disciplinary action by medical boards and prior behavior in medical school. *N Engl J Med.* 2005;353(25):2673-2682.
4. Tamblyn R, Abrahamowicz M, Dauphinee D, et al. Physician scores on a national clinical skills examination as predictors of complaints to medical regulatory authorities. *JAMA.* 2007;298(9):993-1001.
5. Cadieux G, Abrahamowicz M, Dauphinee D, Tamblyn R. Are physicians with better clinical skills on licensing examinations less likely to prescribe antibiotics for viral respiratory infections in ambulatory care settings? *Med Care.* 2011;49(2):156-165.
6. Wenghofer E, Klass D, Abrahamowicz M, et al. Doctor scores on national qualifying examinations predict quality of care in future practice. *Med Educ.* 2009;43(12):1166-1173.
7. Papadakis MA, Hodgson CS, Teherani A, Kohatsu ND. Unprofessional behavior in medical school is associated with subsequent disciplinary action by a state medical board. *Acad Med.* 2004;79(3):244-249.
8. Julian ER. Validity of the Medical College Admission Test for predicting medical school performance. *Acad Med.* 2005;80(10):910-917.
9. Salvatori P. Reliability and validity of admissions tools used to select students for the health professions. *Adv Health Sci Educ Theory Pract.* 2001;6(2):159-175.
10. Kulatunga-Moruzi C, Norman GR. Validity of admissions measures in predicting performance outcomes: a comparison of those who were and were not accepted at McMaster. *Teach Learn Med.* 2002;14(1):43-48.
11. Eva KW, Rosenfeld J, Reiter HI, Norman GR. An admissions OSCE: the multiple mini-interview. *Med Educ.* 2004;38(3):314-326.
12. 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(8):767-775.
13. Reiter HI, Eva KW. Reflecting the relative values of community, faculty, and students in the admis-

- sions tools of medical school. *Teach Learn Med.* 2005; 17(1):4-8.
14. Axelson RD, Kreiter CD. Rater and occasion impacts on the reliability of pre-admission assessments. *Med Educ.* 2009;43(12):1198-1202.
  15. Rosenfeld JM, Reiter HI, Trinh K, Eva KW. A cost efficiency comparison between the multiple mini-interview and traditional admissions interviews. *Adv Health Sci Educ Theory Pract.* 2008;13(1):43-58.
  16. Reiter HI, Eva KW, Rosenfeld J, Norman GR. Multiple mini-interviews predict clerkship and licensing examination performance. *Med Educ.* 2007;41(4):378-384.
  17. Qualifying examination part I reference material and resources. Medical Council of Canada. <http://www.mcc.ca/en/exams/qe1/>. Accessed April 27, 2012.
  18. Qualifying examination part II reference material and resources. Medical Council of Canada. <http://www.mcc.ca/en/exams/qe2/>. Accessed April 27, 2012.
  19. Wood TJ, Humphrey-Murto SM, Norman GR. Standard setting in a small scale OSCE: a comparison of the Modified Borderline-Group Method and the Borderline Regression Method. *Adv Health Sci Educ Theory Pract.* 2006;11(2):115-122.
  20. Lukis K. How to prepare for your dental school interview: tips from an interviewer. <http://wolverinebites.org/2011/03/11/how-to-prepare-for-your-dental-school-interview-tips-from-an-interviewer/>. Accessed October 16, 2012.
  21. Jones PE, Forister JG. A comparison of behavioral and multiple mini-interview formats in physician assistant program admissions. *J Physician Assist Educ.* 2011;22(1):36-40.
  22. Neville AJ, Norman GR. PBL in the undergraduate MD program at McMaster University: three iterations in three decades. *Acad Med.* 2007;82(4):370-374.
  23. Hofmeister M, Lockyer J, Crutcher R. The multiple mini-interview for selection of international medical graduates into family medicine residency education. *Med Educ.* 2009;43(6):573-579.
  24. Norman G. The morality of medical school admissions. *Adv Health Sci Educ Theory Pract.* 2004; 9(2):79-82.
  25. Reibnegger G, Caluba HC, Ithaler D, Manhal S, Neges HM, Smolle J. Progress of medical students after open admission or admission based on knowledge tests. *Med Educ.* 2010;44(2):205-214.
  26. Urlings-Strop LC, Themmen AP, Stijnen T, Splinter TA. Selected medical students achieve better than lottery-admitted students during clerkships. *Med Educ.* 2011;45(10):1032-1040.
  27. O'Neill L, Hartvigsen J, Wallstedt B, Korsholm L, Eika B. Medical school dropout: testing at admission versus selection by highest grades as predictors. *Med Educ.* 2011;45(11):1111-1120.
  28. Puryear JB, Lewis LA. Description of the interview process in selecting students for admission to US medical schools. *J Med Educ.* 1981;56(11):881-885.
  29. Albanese MA, Snow MH, Skochelak SE, Huggett KN, Farrell PM. Assessing personal qualities in medical school admissions. *Acad Med.* 2003;78(3):313-321.
  30. Ambady N, Bernieri F, Richeson J. Toward a history of social behaviour: judgmental accuracy from thin slices of the behavioral stream. In: Zanna MP, ed. *Advances in Experimental Social Psychology*, Vol 32. 2000:201-272.
  31. Dodson M, Crotty B, Prideaux D, Carne R, Ward A, de Leeuw E. The multiple mini-interview: how long is long enough? *Med Educ.* 2009;43(2):168-174.
  32. Ellis APJ, West BJ, Ryan AM, DeShon RP. The use of impression management tactics in structured interviews: a function of question type? *J Appl Psychol.* 2002;87(6):1200-1208.

Education should be a lifelong process, the formal period serving as a foundation on which life's structure may rest and rise.

—Robert H. Jackson (1892-1954)