

Using Multiple-Mini-Interviews for Admission into Teacher Education

Barbara Salingré

University of the Fraser Valley

Sheryl MacMath

University of the Fraser Valley

Abstract

This article reports on the validity of the Multiple-Mini-Interview (MMI) method as it was used by a post-degree teacher education program as part of their admissions process to select candidates for entry into the program. The MMI, primarily used for medical school admittance, involves several stations with different interviewers. Comparisons were made between the MMI, other intake variables, and outtake measures. Quantitative analyses also examined possible interviewer, station, gender, and heritage effects. Results support the claim that the MMI can be used to measure dispositions not measured by other intake variables; however, some concerns did emerge.

Keywords: multiple-mini-interview, teacher education, admission processes, higher education

Résumé

Cet article rend compte de la validité des « mini-entrevues multiples » (MEM), processus d'admission qu'un programme de formation à l'enseignement pour les post-diplômés a utilisé pour la sélection de candidats à l'entrée de leur programme. Ce processus, principalement utilisé pour l'admission aux écoles de médecine, implique plusieurs stations avec différents intervieweurs. Ici, des comparaisons sont faites entre les MEM, d'autres critères d'admission, ainsi que des mesures de sortie. Des analyses quantitatives ont également permis d'examiner les effets discriminatoires possibles provenant de l'intervieweur, de la station, du sexe et de l'origine. Les résultats soutiennent l'affirmation selon laquelle les MEM peuvent être utilisées pour mesurer certaines dispositions que d'autres styles de processus d'admission ne permettent pas ; cependant, certaines préoccupations émergent.

Mots-clés : mini-entrevues multiples, formation à l'enseignement, processus d'admission, études supérieures

Introduction

Our Bachelor of Education program is offered within a small post-degree professional program in a university in southwestern Canada, graduating between 60 and 90 applicants every year. Annually, we receive between 170 and 250 applications and it has become increasingly more competitive for applicants to receive an offer of admission. Our department consists of four full-time faculty members and a number of sessional instructors and non-teaching faculty. We all work toward ensuring that our admission processes and decisions are fair, equitable, defensible, and transparent. We currently use a variety of admission variables: GPA (based on the last 60 course credits completed), program preparation (including work experience with children or adolescents as documented in a résumé) and academic preparation, two references (evaluated as complete/incomplete only), and an interview. Prior to the 2018–2019 cohort year, we utilized a behavioural interview for five years. Faculty members became concerned regarding whether the behavioural interview process itself favoured a specific type of teacher candidate, specifically a female of European heritage with a liberal arts degree and who had spent some time volunteering in classrooms. We were also concerned that interviewers might be influenced in their assessment, both in the interview process itself as well as through having access to the applicant's file, which included their transcript and résumé.

We looked for an alternative to the behavioural interview that would enable us to: (a) evaluate dispositions in line with our departmental values, rather than just classroom experience (which was already assessed as part of their work experience), (b) limit any potential influence the other variables had on the assessment of the interview, and (c) potentially widen access to diverse applicants. After an examination of other professional program admissions, we came across literature and data on the multiple-mini interview (MMI) method, mainly used for medical school admission. This research showed promise to be able to measure non-cognitive dispositions, potentially enabling us to diversify the demographics of admitted students while reducing interviewer bias. We subsequently implemented the MMI in the 2018–2019 admission cycle and committed to researching the feasibility, validity, and acceptability of this admission variable. This article reports on the strengths and challenges of using the MMI in teacher education.

Literature Review

Admission to Teacher Education Programs

Admission decisions to teacher education programs are critical, as there is a high rate of retention and completion of teacher candidates in these programs, thus acting as unofficial gatekeepers to the profession (Thomson et al., 2011). Currently, Canadian teacher education programs utilize a variety of admission variables to determine which applicants are admitted into a program. The most common include GPA, written responses, reference letters, and work experience (Casey & Childs, 2011; Caskey et al., 2001; Crocker & Dibbon, 2008). Far fewer programs utilize a behavioural interview as an admission variable, given the time constraints and costs involved; interviews have been found to require human and financial resources to set up infrastructure and develop the expertise for this admission variable. As 29 out of 40 programs in Canada utilize a combination of written essays, work experience, references, and (in rare cases) interviews (Crocker & Dibbon, 2008), a considerable amount of time is invested in admissions.

Since intake variables for Canadian teacher education programs vary greatly—from GPA, to written responses, to related work experiences, to reference letters—consistent results regarding the value of these admission variables are hard to find (Casey & Childs, 2011; Caskey et al., 2001; Crocker & Dibbon, 2008). Concerns with these variables have arisen regarding validity, evaluator bias, and reliability (Byrnes et al., 2003; Casey & Childs, 2011; Salzman, 1991). As admissions are based on a combination of admission variables, it is important to specifically identify which cognitive or non-cognitive attribute or disposition is measured with each single variable, and if there is correlation amongst the variables used. Hirschhorn and Sears (2015) argue that education program admission needs to move away from a “traditional gatekeeping function” to serve a pedagogical function. This can include the permeation of identified programme values into the criteria and values measured through the admissions process (MacMath & Salingré, 2015).

Conducting individual behavioural interviews (panel of interviewers asking questions that require applicants to reflect on their own experiences) is one of the traditional assessment tools that is time consuming and costly (Casey & Childs, 2011). When

interviews are included as an admission variable, inconsistencies and disagreements among raters are mentioned, even though interviewer training is provided (Childs et al., 2009). In addition, behavioural interviews have been found to have inconsistent inter-rater reliability, making these potentially unfair or unethical (Eva et al., 2004). These concerns are balanced against research that revealed a correlation between interviews and success in practicum (MacMath & Salingré, 2015). However, an additional concern for this small post-degree teacher education program in British Columbia was the ability of our variables to maximize the diversity of successful applicants. While teachers are often Caucasian middle-class females (Childs & Ferguson, 2016), the student population is diverse, creating a need to diversify our teaching force (Childs et al., 2011; DeLuca, 2015). The desire to ensure diversity in applicants being admitted to professional education programs mirrors a similar desire for diversity in admission to health profession programs. Outreach and retention of a diverse population of teacher candidates has been a goal of many education programs to enable equity through admissions (Childs et al., 2011). A review of teacher education literature by Klassen and Kim (2019) suggests that the fairness of current selection methods for underrepresented groups should be investigated and that inter-disciplinary examinations between education and other fields such as medicine may lead to a more robust evidence base. Holden and Kitchen (2016) point to the MMI as utilized in admission to the medical profession as having validity as a non-cognitive assessment tool.

The MMI as One Measure of Admission

Similar to education programs, admission selection to medical schools is considered “high stakes,” as the number of applicants greatly exceeds the number of available spaces in the program (Rees et al., 2016). In addition, admission to medical schools can be comparable to admission to teacher education programs as they are often equivalent to entry into the teaching profession (Thomson et al., 2011; Ferguson et al., 2002). The desire to ensure diversity in applicants being admitted to professional education programs is also similar to health profession programs (Garlick & Brown, 2008; Thomson et al., 2011; Ferguson et al., 2002). There is an equal challenge of widening access and diversity with medical program admission regarding race, ethnicity, or social class, and it is becoming

increasingly important to reflect society within the medical profession (Patterson, Knight et al., 2016) akin to the teaching profession.

The MMI is widely used in professional programs such as medicine, nursing, midwifery, paramedic, and occupational therapy to assess non-cognitive dispositions for admission (Callwood et al., 2018; Knorr & Hissbach, 2014; Patterson, Knight et al., 2016; Pau et al., 2013; Thomas et al., 2017). It is a highly structured applicant selection method designed to assess attributes that best meet the program's requirements with the potential to replace the behavioural interview. Applicants rotate through a series of stations designed to assess non-cognitive attributes. Each station is designed in such a way that specific learned knowledge is not assessed. Rather, an applicant's ability to critically think about a problem and logically work through it while communicating ideas clearly is evaluated (Eva et al., 2004). There is evidence that the MMI can assess an applicant's critical thinking ability, communication skills, professionalism, and attitudes towards certain ethical and social dilemmas (Pau et al., 2013). Unlike a behavioural interview, which focuses on an applicant's reliance on sharing previous experiences in relation to a question, the MMI stations require applicants to respond to a prompt, such as an article, an image, or a quote.

The MMI involves a series of stations with separate interviewers. This allows multiple points of insight into an applicant's abilities and dispositions (Eva et al., 2004). Should an applicant not do well in one station, they have the ability to recover in another station with a new, independent interviewer. Applicants rotate through these stations within a specified period of time, responding to a single prompt at each station (Knorr & Hissbach, 2014; Rees et al., 2016; Pau et al., 2013). The MMI has been reported to provide more valid predictors of inter- and intrapersonal (non-academic) attributes compared to personal statements or references (Patterson et al., 2016). Interviewers receive training before interviewing applicants and are instructed to provide very little feedback to the applicants. One of the other important components of the MMI is that the interviewers are not privy to applicants' background, work experience, or other scores (Jerant et al., 2018). This enables them to use their professional judgement to evaluate the individual's response rather than their overall application, avoiding what Goho and Blackman (2006) describe as a "halo effect," or the tendency to generalize from one area to another.

The separation of all admission components is purposeful, in the hopes that MMI ratings are not influenced by other components. This lack of relationship aims to have

the MMI assess attributes that are different from the other components of the applicant's application (Thomas et al., 2017), enabling greater discriminatory power (Jones & Forster, 2011). Previous studies indicate that the MMI appears to measure something separate from other variables (Eva et al., 2004; Jones & Forster, 2011; Lemay et al., 2007; Patterson, Knight et al., 2016). The MMI has been reported to be a more robust method of selecting medical students with a greater degree of predictive validity (Patterson, Rosetta et al., 2016). In addition, studies report minimal gender, ethnicity, or interviewer effects (Brownell et al., 2007; Griffin & Wilson, 2010; Knorr & Hissbach, 2014; Rees et al., 2016; Uijtdehaage et al., 2011). Many studies into the MMI process have indicated that this method is more objective than other admission variables, and does not appear to be influenced by age or academic level (Callwood et al., 2018). A literature review of 66 publications on the MMI method by Knorr and Hissbach (2014) reported that reliability and feasibility appear to be strengths of the MMI process, but more research is needed with regard to construct validity and comparisons of station types.

Patterson et al. (2016) concluded that the MMI is a more effective selection method, which tends to be generally fairer than behavioural interviews, references, and personal statements. In addition, it appears to enable a wider sampling of applicant competencies, which may provide programs with a better analysis of their overall ability (Pau et al., 2013). As recommended by Holden and Kitchen (2016), the MMI presents a potential variable to address a number of these concerns for teacher education admissions.

The MMI in Teacher Education Admission

Before using the MMI with our post-degree program, we examined results, even though they originated from medical programs, in relation to whether the MMI was feasible for us to implement. Brownell et al. (2007) reported that using the MMI for candidate selection does not require more examiners than behavioural interviews, which are often conducted in a panel format. It also did not necessarily incur a greater cost, as interviews could be completed efficiently over a short period of time. In comparison to behavioural interviews, the MMI has been found to be as cost-efficient as other interview formats (Rosenfeld et al., 2008). While investments in the first year of designing and implementing an interview infrastructure were greater, the process became more efficient and feasible over time for the size of the program and the number of applicants.

The MMI has been perceived as fair and transparent by both interviewers and interviewees (Uijtdehaage et al., 2011), thus leading to positive experiences for everyone involved (Eva et al., 2004). In addition, there are further research results that show that, in the case of medical students, successful applicants end up preferring the schools that conducted the interviews (McManus et al., 1999; Patterson et al., 2016). There is also evidence that the MMI interview, which provides a great deal of structure, is preferred over less-structured methods (Eva et al., 2004). This is further underscored by Patterson et al. (2016), who stated that there is tentative evidence that MMIs and more structured interviews are preferred by applicants. It is noteworthy that information provided to applicants about the MMI beforehand is necessary. Timing considerations (number of stations and length), stress on applicants, and interview style preferences matter to applicants (Rees et al., 2016). The overall experience of the MMI has been found to be positive for both applicants and interviewers (Eva et al., 2004). Whether or not this level of acceptability of the MMI as one of the admission variables to health care programs would be the same for a post-degree teacher education program is something that needs to be investigated over time.

Although traditional interviews have been used as part of teacher education admission (Byrnes et al., 2003; Casey & Childs, 2011; Childs et al., 2009; Crocker & Dibbon, 2008; Harrison et al., 2012), the MMI has not been used, even though non-cognitive dispositions are routinely measured as part of the admissions process to this professional program. The examination of the MMI specific to a Canadian post-degree teacher education program was undertaken to investigate three factors: (a) can this variable assess non-academic dispositions not measured by other admission variables, (b) do station and/or interviewer effects have an influence on the scoring of applicants, and (c) is there a correlation between the MMI and success in long practicum (outtake measures)?

Methodology

This is a quantitative study designed to evaluate correlations between and amongst admission variables and outtake variables (faculty mentor ratings of success in long practicum) over one admission cycle to a 10-month post-degree teacher education program. The variables used for admissions include an interview, evaluation of program

preparation (e.g., time in K–12 classrooms, academic preparedness, international experience, etc.), and GPA (a minimum of 2.8 on a 4.33 scale is required). These admission scores were assessed and compared using correlation and ANOVA tests of significance to determine those admission variables that appear to measure a variety of applicant skills, dispositions, and experiences.

Within the admission cycle for the 2018–2019 cohort, 20% of the admission score was assigned to the GPA, 40% to program preparation (including the assessment of previous work experiences with children in an educational setting as demonstrated through the submission of a résumé and academic preparation), and 40% of the admission score was derived from the MMI.

Data Sources

Given the small size of our teacher education program, it was viable for us to use the MMI for all eligible applicants to the 2018–2019 cohort. Three months prior to the admission deadline, we struck an Admissions Committee who reviewed all aspects of the admission procedures. This committee reviewed the literature on the MMI and decided that it would be feasible to start with four stations. We began to develop prompts/questions for the MMI reflecting the department values using the dispositions of critical thinking, communication, compassion, and professionalism. Questions for each station were to include a variety of prompts, such as scenarios, images, audio recordings, articles, or a hands-on opportunity. Each member of the Admissions Committee was then assigned one of the questions/prompts for further development, including background information and follow-up questions. These prompts were peer-reviewed and finalized six weeks later. We then began recruiting teachers and administrators within the local school districts to act as interviewers. Faculty Mentors (FMs) who currently supervise teacher candidates in the schools during their practicum were also recruited as interviewers. The criteria for volunteer interviewers was that they had to be educators from the K–12 field, as we wanted their professional appraisal of each applicant's performance in the interview. We then developed a two-hour training session for the interviewers with demonstrations so that they felt comfortable scoring applicants on the four specific dispositions. The scoring sheet listed these four questions with a Likert scale from one to five: Was the applicant able to reflect critically? Did the applicant provide and articulate a thoughtful response?

Was the applicant able to demonstrate compassion for a multiplicity of views? Was the demeanour and conduct of the applicant professional? We also requested that interviewers sign a disclosure agreement protecting the confidentiality of the questions. In addition, we requested that interviewers report any pre-existing connections with any of the applicants, to avoid any intended or unintended perception of bias.

MMI station descriptions. Interviews were scheduled on weekends, which allowed greater access to interviewers in the field and adjoining rooms to house each station. At the beginning of each interview cycle, applicants were introduced to the process, and the structure and timing for each station was explained. Applicants were asked to sign a nondisclosure form, so questions would not be shared amongst future applicants. They were also told that interviewers may display a more neutral affect, to avoid the applicant seeking affirmation from the interviewer during the conversation. The template for each station was the same. At every station, the student would have two minutes to read the prompt on the outside of a door. The applicant could collect their thoughts and, at the sound of a bell, would enter the room and meet the interviewer. The applicant would then have seven minutes to respond to the prompt. The interviewer was instructed not to interrupt the commentary, but was provided with background information for the prompt, along with follow-up questions should the applicant freeze or get stuck due to personal anxiety or confusion. It is important to note that the interviewer had no other information about the applicant and only scored the applicant after they left the room. Below is an example from one of the stations:

You are a teacher getting ready for a meeting with one of your student's parents. The parent is coming to discuss how your student is transitioning from female to male. How do you feel about this? What concerns do you have?

Interviewers were provided with follow up questions, if needed. These were as follows:

1. How you will prepare for this meeting?
2. What key points will you want to include in this meeting?
3. What are the different perspectives relevant to this scenario?
4. What is your responsibility in this scenario?
5. How does this scenario connect to your understanding of schools?

At the end of the whole interview cycle, each applicant would have 16 scores (four from each interviewer), which were then averaged/combined to form a single score for the interview.

The four stations for the 2018–2019 intake were anchored by a newspaper article, a scenario (provided as an example above), a set of images, and a quote. We interviewed 156 applicants with four stations and 14 interviewers. Out of the pool of 156 students who interviewed for admission, 94 received an offer of admission. For those 94 successful applicants, quantitative analyses were completed to examine correlations between admission variables (program preparation, interview, and GPA), as well as possible interviewer or station effects. Both gender and applicant ethnicity were analyzed for possible effects. These admission scores were then correlated with an outcome measure completed at the end of the program by the FMs responsible for observing (minimum of 17 visits, one per week) and evaluating teacher candidates. The outcome measure utilized a five-point Likert scale (1 = poor, 5 = exceptional) with FMs evaluating each teacher candidate on their effectiveness in relation to 14 measures (e.g., classroom management, effective teaching practices, connecting with students, valid and reliable assessment practices, etc.).

Results

For the 2018–2019 cohort year there were 156 completed applications and, after the admissions process, 94 applicants were offered admission (75 females, 19 males; 64 elementary, 20 secondary humanities, and 12 secondary math/science). The majority of accepted applicants were of European heritage, while 10 identified as having Indigenous heritage and another 10 identified as having South Asian heritage. Analyses were completed only for those candidates who were admitted to the program and consented to have their data included in this research. It did not seem ethical to request permission of applicants prior to the interviews to have their data included in this study. This exclusion represents a limitation of this analysis. The descriptive statistics for the 2018–2019 cohort year can be found in Table 1.

Table 1*Intake Variable Means and Standard Deviations*

2018–2019	<i>N</i> = 94	
	<i>M</i>	<i>SD</i>
GPA (out of 2)	1.46	.61
Program prep (out of 4)	3.21	.55
MMI (out of 4)	2.87	.47
Total intake score (out of 10)	7.60	1.07

Given the time commitment to develop the stations, recruit and train the interviewers, and organize the interview sessions, the Admissions Committee estimated that the same number of department hours were spent on this process, as compared to previous admission cycles where behavioural panel interviews were used. This was due mainly to the use of FMs and external administrators and teachers from the field; in previous years the behavioural interviews were conducted by department faculty only.

Looking for Potential Discrimination

As our study was initiated with the hope of ensuring that a diverse set of skills, dispositions, and experiences were considered as criteria for admission, it was important to investigate whether any of the admission variables discriminated against applicants. We looked at gender, heritage, and program selection as possible areas of discrimination. We used ANOVA tests on each intake variable and, if the Levene's test of homogeneity of variance was significant, a Chi-square non-parametric test was used. For the 2018–2019 year, there were no differences due to gender in relation to any intake variable: GPA, $F(1, 92) = .00, p = .96$; overall MMI score, $F(1, 92) = .33, p = .57$; program preparation, $F(1, 92) = .10, p = .75$; or total scores, $F(1, 92) = .00, p = 1.00$. There were also no differences by program in relation to GPA, $F(2, 91) = 2.08, p = .13$; program preparation, $F(2, 91) = 2.10, p = .13$; or total scores, $F(2, 91) = 1.03, p = .36$. As the Levene's test was significant for an analysis of the overall MMI score, the Chi-square revealed no significant differences due to program in relation to the overall MMI score, $X^2(68, N = 94) = 71.06, p = .38$.

In relation to heritage, there were no significant differences for GPA, $F(2, 91) = 1.73, p = .18$. However, according to the Chi-square non-parametric test there were significant differences in relation to program preparation, $X^2(18, N = 93) = 41.15, p = .00$; candidates reporting European heritage ($M = 3.29, SD = .47$) scored significantly higher than both groups of candidates reporting South Asian ($M = 2.90, SD = .54$) or Indigenous ($M = 2.90, SD = .87$) heritage. As well, there were significant differences in relation to the overall MMI score, $F(2, 91) = 4.17, p = .02$. A Tukey HSD comparison revealed that candidates reporting South Asian heritage ($N = 10, M = 7.79, SD = .97$) scored significantly higher than candidates reporting Indigenous heritage ($N = 10, M = 6.71, SD = 1.40$). To examine this further, we looked at specific stations being used to generate the overall MMI score. Only the news article station showed significant differences, $F(2, 91) = 4.56, p = .01$; a Tukey HSD comparison revealed that candidates reporting European ($N = 74, M = 14.41, SD = 3.40$) and South Asian heritage ($N = 10, M = 16.20, SD = 4.24$) scored significantly higher than candidates reporting Indigenous heritage ($N = 10, M = 11.50, SD = 4.04$). Finally, there were significant differences in heritage and the total overall score, $F(2, 90) = 9.26, p = .02$; a Tukey HSD comparison was calculated and candidates reporting European heritage ($N = 74, M = 7.70, SD = .98$) or South Asian heritage ($N = 10, M = 7.79, SD = 1.06$) scored significantly higher than candidates reporting Indigenous heritage ($N = 10, M = 6.71, SD = 1.40$). These results are concerning; however, a great deal of caution is needed in their interpretation given the small sample sizes for those reporting Indigenous or South Asian heritage.

MMI Station and Interviewer Effects

For the 2018–2019 year, we explored possible effects in relation to the different MMI stations and interviewer types. There were no station effects, $F(3, 1484) = 2.087, p = .10$. However, a Chi-square test revealed that external evaluators ($N = 276, M = 3.98, SD = .93$) scored significantly higher than both faculty ($N = 756, M = 3.49, SD = 1.15$) and FMs ($N = 456, M = 3.55, SD = 1.09$), $X^2(8, 1488) = 47.637, p = .00$. To examine these differences further we utilized an ANCOVA [between-subjects factor: type of interviewer (faculty, FMs, external); covariate: station (images, news article, scenario, quote)] and revealed an interaction, $F(5, 1477) = 6.35, p = .00$. As external evaluators did not

score the images station, faculty and FMs were significantly different, $F(1, 370) = 8.97$, $p = .00$. See Tables 2 and 3 for a summary of the ANCOVA and Bonferroni Comparisons.

Table 2

Analysis of Covariance Summary: 2018–2019 Interviewer and Station Effects

Source	SS	df	MS	F	p
Station	9.92	3	3.31	2.80	.04
Interviewer Type	48.73	2	24.37	20.65	.00
Station*Interviewer	31.73	5	6.35	5.38	.00
Error	1742.81	1477	1.18		

Table 3

Bonferroni Comparisons for Type of Interviewer by Station

Comparisons				95% CI			
Station	Interviewer Type	MD	SE	p	Lower Bound	Upper Bound	
News Article	Faculty	FMs	-.25	.13	.16	-.55	.06
		External	-.38 ^a	.14	.02	-.72	-.05
	FMs	External	-.14	.13	1.00	-.49	.21
Scenario	Faculty	FMs	.36 ^a	.14	.03	.03	.68
		External	-.22	.15	.44	-.57	.14
	FMs	External	-.57 ^b	.15	.00	-.92	-.22
Quote	Faculty	FMs	-.02	.12	1.00	-.31	.27
		External	-.77 ^b	.14	.00	-1.10	-.43
	FMs	External	-.75 ^b	.14	.00	-1.09	-.40

^a Difference significant at the .05 level.

^b Difference significant at the .00 level.

This extended analysis reveals that while the scores between faculty and FMs overall did not differ significantly, when broken down by stations, there were significant differences. While faculty ($M = 3.43$, $SD = .06$) scored significantly lower than the FMs ($M = 3.93$, $SD = .15$) at the images station, faculty ($M = 3.71$, $SD = .10$) scored significantly higher than the FMs ($M = 3.35$, $SD = .09$) at the scenario station.

Correlations Between Intake Variables

Beyond trying to avoid possible discriminatory effects in our attempt to diversify the characteristics we are considering during the admissions process, we looked at the correlations between intake variables. It was our goal that there would be no correlations amongst the intake variables, indicating that they may be measuring different characteristics. Using the Pearson Correlation Coefficient, we compared the three intake variables and the total overall score. Calculations reveal significant correlations between GPA ($M = 1.46$, $SD = .61$) and the MMI ($M = 2.87$, $SD = .47$), and no significant correlations with program preparation ($M = 3.21$, $SD = .55$). See Table 4 for results.

Table 4

Intercorrelations between Intake Variables

Year	Variables	1	2	3	4
2018–2019	GPA	—	.34 ^a	.05	.69 ^a
	MMI		—	.06	.66 ^a
	Program prep			—	.48 ^a
	Total				—

^a Correlation is significant at the 0.01 level (2-tailed).

Correlations with Outtake Measures

Each of the intake measures were compared with FM scores after completion of the long practicum on the 14 program goals. FMs had been using the program goals throughout the 17 weeks of practicum experiences and were very familiar with them. Results are detailed in Table 5. Unfortunately, very few significant correlations were revealed between any of the intake variables. These results suggest that the intake variables, including the MMI, did not correlate with those considered most successful in the program.

Table 5*Correlations between Intake Variables and Outcome Variables*

Intake*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Total
GPA	.01	.06	.08	.07	.05	.02	.07	.18	.06	.13	.12	.14	.04	.06	.09
MMI	.08	.19	.15	.15	.21 ^a	.15	.17	.14	.21 ^a	.18	.11	.20	-.08	.09	.17
Program prep	-.13	-.05	-.02	-.19	-.09	.04	.04	-.06	-.00	-.16	-.10	.06	-.02	-.06	-.07
Total score	-.09	.06	.06	-.06	.03	.07	.09	.04	.09	.03	.01	.12	-.07	.01	.03

*1 = professionalism, 2 = connecting with students, 3 = classroom management, 4 = reflective practice, 5 = working with feedback, 6 = clear and observable vision, 7 = knowledge of students, 8 = knowledge of content, 9 = effective communication, 10 = effective teaching practices, 11 = reliable assessment, 12 = inclusive and individualized, 13 = working with guardians, 14 = contribute to school community.

^a Correlation is significant at the 0.05 level (2-tailed).

Discussion

Three research questions motivated this study: (a) can the MMI assess non-academic dispositions not measured by other admission variables, (b) do station and/or interviewer effects during MMIs have an influence on the scoring of applicants, and (c) is there a correlation between the MMI and success in long practicum (outtake measures)? We examine each of these questions in turn.

Can This Variable Assess Non-Academic Dispositions Not Measured by Other Admission Variables?

Similar to medical school studies examining the ability of the MMI to evaluate dispositions in medical students (Callwood et al., 2018; Knorr & Hissbach, 2014; Patterson et al., 2016; Pau et al., 2013; Thomas et al., 2017), our results indicate a lack of

correlation on scores of program preparation and the MMI. This was important to us, as our previous behavioural interviews appeared to assess work experience as opposed to dispositions. In contrast, the MMI did not appear to measure responses that were influenced by work experiences. We were very encouraged by this finding, as the goal of utilizing the MMI was to measure something different from what appeared on applicants' résumés. Our assumption is that the MMI enabled us to measure actual dispositions. Additional studies are needed to confirm this. We did find it surprising that there was a significant correlation between GPA scores and the MMI. We are not clear on why this is the case. This may have occurred because students with a higher GPA have demonstrated greater levels of critical thinking in their university course work. Again, more investigation is needed into this correlation.

We feel that it is important to note that, while the MMI did appear to enable us to consider a wide variety of candidates' skills, experiences, and dispositions, this does not guarantee a more diverse teaching force. When looking at our statistics as to who makes it into our program, we still have a majority of young females with European heritage in the applicant pool. Childs et al. (2011) explore the idea of equity in and equity through admissions. Equity in admissions focuses on the fact that everyone should have equal access and consideration during the admission process. In contrast, equity through admissions is about having admission processes that enable the selection of those teacher candidates with a specific social justice and equity disposition. With equity through admission, it is not about having, for example, greater ethnicities in our teaching force, but having a teaching force that teaches equitably. We do not have data if our admission processes accomplish this, but it is something we wish to examine further.

Given our concern with equity, it is worth noting that our results revealed no gender effects. In contrast, the MMI did appear to favour those with European heritage, though these findings require extreme caution, given the limited ethnic diversity of our applicants. While our program reserves 10% of its seats to applicants with Indigenous heritage, we rarely have enough applicants to meet this specific quota.

Do Station and/or Interviewer Effects Have an Influence on the Scoring of Applicants?

Our results in this area need to be taken with extreme caution. Knorr and Hissbach (2014) recommend a minimum of 6 to 12 stations, while Rees et al., (2016) state that 7 to 12 stations are optimal. Due to limited faculty and time resources, we were only able to utilize four stations, and this could impact our results. Our findings revealed no overall station effects. In contrast, we did find interviewer effects. External evaluators who were not as familiar with our program values scored applicants significantly higher in every station. While faculty mentor and faculty scores appeared to correlate, an ANCOVA revealed significant differences based on station. Faculty mentors scored significantly higher at the images station, while faculty scored significantly higher in the scenario station. The images station required a critical analysis of the education system while the scenario focused on solving a practical problem common in schools. We postulate that faculty mentors whose experience is based in schools found it easier to be critical of the scenario, as they were most familiar with it. The images station, which required a critical analysis of the education system that they worked with, may have been more challenging. As a result, faculty mentors scored applicants higher. This hypothesis requires further examination.

Regardless of the reasons, we are concerned that interviewer effects were revealed. Our findings contrast with the majority of results in the medical field where interviewers are also internal and external to the program (Knorr & Hissbach, 2014; Rees et al., 2016; Thomas et al., 2017). There are many possible reasons for this: too few stations, lack of adequate interviewer training, or the diversity in the type of interviewer recruited. Implementing the MMI as part of our admission processes was only feasible (in contrast to a behavioural interview) because only four stations were developed and we were able to use external interviewers. The Admissions Committee for our program believes that it is worthwhile to continue with the MMI, though altering the training for the interviewers and the type of stations may be necessary. However, if these effects cannot be reduced, it may mean that the MMI may not be a reasonable alternative to the behavioural interview in the long term.

Is There a Correlation Between the MMI and Success in Long Practicum (Outtake Measures)?

The MMI only correlated with two of our outtake measures: working with feedback and effective communication. To us, the correlation between effective communication and the MMI results is not surprising, as one of the criteria that the evaluators assessed was, in fact, communication. We are concerned that there were no significant correlations between the MMI and the overall outtake score. There are many possible reasons for this lack of significant correlation beyond the effectiveness of the MMI. This was the first year in which we admitted 96 candidates in the program (we typically have 64). Again, we only had four stations as opposed to the seven that are recommended for optimal admission processes. As a result, more data and investigation are required.

Examining correlations between the MMI and any outtake measures is challenging. In the medical field, Kelly et al. (2014) found that the MMI was not predictive of success in the program. Jerant et al. (2018) did report that the MMI significantly correlated with success after a one-year clerkship. There are, however, no other studies examining correlations between MMI scores and outtake measures (Knorr & Hissbach, 2014). In addition, in teacher education there is little longitudinal research examining admission and teacher effectiveness (Klassen & Kim, 2019). Casey and Childs (2011) even question the ability to have reliable outtake measures for a teacher education program. We still think it is valuable to draw comparisons between all admission variables (intake measures) and faculty mentor ratings of teacher performance (outtake measures), however, a lack of significant correlation will not deter us from using the MMI unless future research results indicate a need for change. Our purpose for testing the MMI focused on increasing the diversity of applicant profiles in admissions, not predicting success in long practicum. As such, these results encourage further investigating and monitoring.

Educational Importance of the Study

Our post-degree teacher education program has been the first to utilize the MMI as part of its admission variables. We would recommend that other education programs consider the MMI as an admission variable when trying to diversify the applicant characteristics, dispositions, and experiences being considered. We caution education programs on two

counts: The number of stations is important to consider, as well as interviewer effects. We wonder if the interviewer effects would have been decreased if we had used more stations? It is clear that consideration needs to be given regarding who the interviewers are. While our faculty and faculty mentors had more extensive knowledge about program values, they were more consistent in their scoring. External field evaluators who were not as familiar with our program values tended to mark easier. We would recommend that any other teacher education program that is considering using the MMI as an admission variable invest time and adequate resources in deciding which outtake measures to employ in order to assess the appropriateness of the MMI to predict success in teacher training programs.

References

- Brownell, K., Lockyer, J., Collin, T., & Lemay, J-F. (2007). Introduction of the multiple-mini-interview into the admissions process at the University of Calgary: Acceptability and feasibility. *Medical Teacher*, 29(4), 394–396.
- Byrnes, D. A., Kiger, G., & Shechtman, Z. (2003). Evaluating the use of group interviews to select students into teacher-education programs. *Journal of Teacher Education*, 54(2), 163–172.
- Callwood, A., Cooke, D., Bolger, S., Lemanska, A., & Allan, H. (2018). The reliability and validity of multiple mini interviews (MMIs) in values-based recruitment to nursing, midwifery and paramedic practice programmes: Findings from an evaluation study. *International Journal of Nursing Studies*, 77, 138–144.
- Casey, M., & Childs, R. (2011). Teacher education admission criteria as measure of preparedness for teaching. *Canadian Journal of Education*, 34(2), 3–20.
- Caskey, M., Peterson, K., & Temple, J. (2001). Complex admissions selection procedures for graduate pre-service teacher education program. *Teacher Education Quarterly*, 28(4), 7–21.
- Childs, R. A., Broad, K., Gallagher-Mackay, K., Sher, Y., Escayg, K-A., & McGrath, C. (2011). Pursuing equity in and through teacher education program admissions. *Education Policy Analysis Archives*, 19(24), 1–22.
- Childs, R. A., Ferguson, A. K., Herbert, M. B., Broad, K., & Zhang, J. (2016). Evaluating admission criteria effects for under-represented groups. *Higher Education Research & Development*, 1–11.
- Childs, R. A., Ram, A., & Xu, Y. (2009). Combining dual scaling with semi-structured interviews to interpret rating differences. *Practical Assessment, Research, and Evaluation*, 14(11), 1–9.
- Crocker, R., & Dibbon, D. (2008). *Teacher education in Canada: A baseline study*. Society for the Advancement of Excellence in Education.
- DeLuca, C. (2015). Rethinking the evaluation of university admission policy and practice: A Canadian neoinstitutional perspective. In V. Stead (Ed.), *International*

- perspectives on higher education admission policy: A reader* (pp. 268–279). Peter Lang.
- Eva, K. W., Rosenfeld, J., Reiter, H. I., & Norman, G. R. (2004). An admissions OSCE: The multiple-mini-interview. *Medical Education*, 38(3), 314–326.
- Ferguson, E., James, D., & Madeley, L. (2002). Factors associated with success in medical school: Systematic review of the literature. *British Medical Journal*, 324, 952–957.
- Garlick, P., & Brown G. (2008). Widening participation in medicine. *British Medical Journal*, 336, 1111–1113.
- Goho, J., & Blackman, A. (2006). The effectiveness of academic admission interviews: An exploratory meta-analysis. *Medical Teacher*, 28(4), 335–340.
- Griffin, B. N., & Wilson, I. G. (2010). Interviewer bias in medical student selection. *Medical Education*, 193(6), 343–346.
- Harrison, J., Smithey, G., McAfee, H., & Weiner, C. (2012) Assessing candidate disposition for admission into teacher education: Can just anyone teach? *Action in Teacher Education*, 27(4), 72–80.
- Hirschhorn, M., & Sears, A. (2015). More than gatekeeping: The pedagogical potential of admissions procedures for teacher education. In L. Thomas & M. Hirschhorn (Eds.), *Change and progress in Canadian teacher education: Research on recent innovations in teacher preparation in Canada* (pp. 441–463). Canadian Association for Teacher Education.
- Holden, M., & Kitchen, J. (2016). Evolving practices: Admission policies in Ontario teacher education programs. *Canadian Journal of Education*, 39(4), 1–28.
- Jerant, A., Henderson, M. C., Griffin, E., Hall, T. R., Kelly, C. J., Peterson, E. M., Wofsy, D., & Franks, P. (2018). Do multiple mini-interview and traditional interview scores differ in their associations with acceptance offers within and across five California medical schools? *Academic Medicine*, 93(8), 1227–1233.
- Jones, P. E., & Forster, J. G. (2011). A comparison of behavioral and multiple-mini-interview forms in physician assistant program admission. *The Journal of Physical Assistant Education*, 22(1), 36–41.

- Kelly, M. E., Dowell, J., Husbands, A., Newell, J., O'Flynn, S., Kropmans, T., Dunne, F. P., & Murphy, A. W. (2014). The fairness, predictive validity and acceptability of multiple mini interview in an internationally diverse student population — a mixed methods study. *BMC Medical Education, 14*, 267.
- Klassen, R., & Kim, L. (2019). Selecting teachers and prospective teachers: A meta-analysis. *Educational Research Review, 26*, 32–51.
- Klassen, R., Kim, L., Rushby, J., & Bardach, L. (2020). Can we improve how we screen applicants for initial teacher education? *Teaching and Teacher Education, 87*, 2–11.
- Knorr, M., & Hissbach, J. (2014). Multiple-mini-interviews: Same concept, different approaches. *Medical Education, 48*(12), 1157–1175.
- Lemay, J-F., Lockyer, J. M., Collin, V. T., & Brownell, A. K. W. (2007). Assessment of non-cognitive traits through the admissions multiple-mini-interview. *Medical Education, 41*(6), 573–579.
- MacMath, S., & Salingré, B. (2015). Evaluating intake variables for a teacher education programme: Improving student success and process efficiency. *Studies in Educational Evaluation, 47*(1), 76–82
- McManus, I., Richards, P., & Winder, B. (1999). Do UK medical school applicants prefer interviewing to non-interviewing schools? *Advanced Health Science Education Theory and Practice, 4*, 155–165.
- Patterson, F., Knight, A., Dowell, J., Nicholson, S., Cousans, F., & Cleland, J. (2016). How effective are selection methods in medical education? A systematic review. *Medical Education, 50*(1), 36–60.
- Patterson, F., Rosetta, E., Hale, R., Grant, M., Roberts, C., Cousans, F., & Martin, S. (2016). The predictive validity of a situational judgement test and multiple-mini interview for entry into postgraduate training in Australia. *Medical Education, 16*, 87–95.
- Pau, A., Jeevaratnam, K., Chen, Y., Fall, A., Khoo, C., & Nadarajah, D. (2013). The multiple mini-interview (MMI) for student selection in health professions training: A systematic review. *Medical Teacher, 35*(12), 1027–1041.

- Rees, E., Hawarden, A., Dent, G., Hays, R., Bates, J., & Hassell, A. (2016). Evidence regarding the utility of multiple mini-interview (MMI) for selection to undergraduate health programs: A BME systematic review. *Medical Teacher, 38*(5), 443–445.
- Rosenfeld, J., Reiter, H., Trinh, K., & Eva, K. (2008). A cost efficiency comparison between the multiple mini-interview and traditional admission interviews. *Advances in Health Science Education, 13*(1), 43–58.
- Salzman, S. (1991, February 16–20). *Selecting the qualified: Predictors of student teacher performance* [Paper presentation]. Annual Meeting of the Association of Teacher Educators, New Orleans, LA.
- Thomas, A., Young, M. E., Mazer, B. L., Lubarsky, S., & Razack, S. (2017). Reliability and validity of the multiple mini-interview (MMI) for admissions to an occupational therapy professional program. *British Journal of Occupational Therapy, 80*(9), 558–567.
- Thomson, D., Cummings, E., Ferguson, A. K., Miyuki Moizumi, E., Sher, Y., Wang, X., & Childs, R. A. (2011). A role for research in initial teacher education admissions: A case study from one Canadian university. *Canadian Journal of Educational Administration and Policy, 121*, 1–23.
- Uijtdehaage, S., Doyle, L., & Parker, N. (2011). Enhancing the reliability of the multiple-mini-interview for selecting prospective health care leaders. *Academic Medicine, 86*(8), 1032–1039.