



Multiple Mini-interview Consistency and Satisfactoriness for Residency Program Recruitment: Oman Evidence

Rashid Al Abri¹, John Mathew¹ and Lakshmanan Jeyaseelan²

¹Ear, Nose and Throat Division, Department of Surgery, College of Medicine and Health Sciences, Sultan Qaboos University, Muscat, Oman

²Department of Biostatistics, Christian Medical College, Vellore, India

ARTICLE INFO

Article history:

Received: 23 August 2017

Accepted: 19 December 2018

Online:

DOI 10.5001/omj.2019.42

Keywords:

Education; Interview.

ABSTRACT

Objectives: Standard interviews are used by most residency programs to assess non-cognitive skills, but variability in the interviewer's skills, interviewer bias, and context specificity limit reliability. We sought to investigate the consistency and satisfactoriness of the multiple mini-interview (MMI) model for resident selection into an otorhinolaryngology head and neck surgery residency program. **Methods:** This pilot study was done in an independent academic residency training center for 15 applicants, in seven eight-minute MMI stations with eight raters for the 2015–2016 academic year. The raters included the chief resident and education committee chairman in one of the stations. Candidates were assessed on two items: medical knowledge (two standardized case scenarios) and behavioral knowledge (personality and attitude, professionalism, communication, enthusiasm to the specialty, and English proficiency). **Results:** Of 15 candidates, 10 (66.7%) were female and five (33.3%) were male; five were recommended for selection, and five were kept on the waiting list. The reliability, intraclass correlation coefficient (ICC), of the scores obtained from seven items of MMI was 0.36 (95% confidence interval (CI): -0.31–0.75; $p = 0.110$). However, the ICC of the medical interview was 0.54 (95% CI: 0.45–0.84; $p = 0.090$). The correlation between behavioral items score and MMI total score was $r = 0.135$ ($p = 0.150$). **Conclusions:** The interview evaluation/survey form given to candidates and interviewers has shown that MMI is a fair and effective tool to evaluate non-cognitive traits. Both candidates and interviewers prefer MMI to standard interviews. The MMI process for residency interviews can generate reliable interview results using only seven stations and is acceptable and preferred over standard interview modalities by residency program applicants and faculty members.

The admission procedure in medical schools is the paramount assessment exercise. There is considerable controversy too due to different methods of evaluation in practice. Besides academic excellence, health science programs value non-intellectual variables such as social or relational skills, honesty, and clinical acumen.¹ Six domain of competencies to select candidates for training such as patient care, medical knowledge, system based practice, practice-based learning, clinical acumen, and social or relational skills were recommended by the Accreditation Council for Graduate Medical Education (ACGME) in 1999.² The multiple mini-interview (MMI) model was first developed in 2001 to reduce interviewer bias and increase the accuracy of candidate scoring by expanding the number of interviewers and standardizing interview questions.¹

Various medical schools and recruitment programs for residents in Canada and the UK have started using the MMI method since.^{3–5} Evidence for its high reliability has been demonstrated using 6–12 interview stations.^{1,6–8} Interviewers and applicants have found it to be an acceptable alternative to the traditional interview.^{4,5,9} The predictive ability of MMI methods concerning patient care measures and licensing examination score were exhibited only in a few studies.^{10,11} The studies that have used MMI methodology have shown strong statistical reliability and generalizability coefficients.^{10,12,13}

The MMI method has become popular to select residents who could handle stress, make quick and ethically-bound decisions, and work well as a team member. While a single interview would seldom find a person suitable for a resident program, the MMI method is well suited to identify such a person for a

specialty. In a traditional interview, the residents face many committee members at once, and therefore there is an opportunity for a biased selection. There is no opportunity to appear again. However, the MMI method provides multiple opportunities and, therefore, selection due to chance is reduced.

We started to apply these competencies in our Otorhinolaryngology Head Neck Surgery (ORLHNS) residency program, as part of the Oman Medical Specialty Board's ACGME-I accreditation application. In Oman, traditional interview methods are used in the residency programs to grade aptitude for non-intellectual ability to assess for thoroughness in clinical acumen, interpersonal relationships, and communication skills. However, the reliability of this method is questionable due to substantial variability across examiners in scoring candidates and potentially biased view or approach towards a person.^{3,14-21} The process has been described as an "elaborate, labor-intensive lottery."²² Therefore, the aim of the study was to examine the consistency and the satisfactoriness of a seven-station MMI method in the selection of residents in an ORLHNS residency program in Oman with the assumption that the medical assessment score is independent of behavioral assessment score.

METHODS

This pilot study was done in an independent academic residency training center, with approval from the Otorhinolaryngology Education Committee, Oman Medical Specialty Board, for 15 applicants in the 2015–2016 academic year. There were eight interviewers, including the chief resident and education committee chairman in one station.

The seven stations were set in independent rooms, which comprised of different scenarios that assessed the medical knowledge and behavioral skills of the applicants. There were two standard case scenarios to evaluate medical knowledge and five scenarios to measure behavioral knowledge or non-cognitive skills (personality and attitude, thoroughness in clinical acumen, writing and disseminating, passion to the specialty, and English expertise).

On the day of the interview, every applicant completed seven eight-minute MMI stations. The applicants' file or applications were sent to the interviewers two days before the interview. At each station, candidates had two minutes to read

Box 1: Multiple mini-interview and acceptability survey components.

- Multiple mini-interview
- Seven stations
- Eight raters
- Eight minutes each (two minutes to read the instructions or scenario, and six minutes to answer)
- Thirty seconds to move from one station to the next
- Global performance scoring system
- Acceptability survey
- Paper-based
- Self-administered
- Closed-ended
- Five-point Likert's scale scoring system (1 = strongly disagree to 5 = strongly agree).

the instruction or background information on the scenario and six minutes to address the scenario with the interviewer. Each interviewee spent a total of 56 minutes on the entire interview process.

A separate room was allotted for each station, and a single interviewer was assigned to each station except for station four (enthusiasm to the specialty), which had two interviewers: the chairman and chief resident. Candidates were rated per given scenario at each station, and scoring was based on the standard scoring system set by Oman Medical Specialty Board, which was provided in the applicant interview evaluation form. Interviewers could make notes on the marking sheet regarding any problems or issues that may have come up during the interview.

A summary of the layout of the MMI and acceptability survey are given in Box 1.

Applicants were asked to fill the survey form anonymously based on their observations regarding the MMI in the context of impartiality, stress level, and efficiency in assessing their non-intellectual traits. Interviewers who performed MMI were also asked to fill the survey form in assessing the usefulness of MMI process in the domains of being impartial, efficient in evaluating the non-intellectual skills, and their preference compared to a standard interview process.

Excel was used to enter the data from the interview for the computation and ranking of the candidates. Each domain score was described by the mean, standard deviation (SD), minimum and maximum. Scatter plots and Pearson's correlation coefficients were used to present the relationship between two domains score. The reliability of the seven station scores were assessed using intraclass correlation

Table 1: Mean values of study variables in 15 residents interviewed by the multiple mini-interview model.

Study variables	Mean	SD	Minimum	Maximum
Age, years	25.8	2.2	23.0	30.0
International format examination score	457.2	50.6	402.0	577.0
Medical knowledge	4.3	1.2	3.0	6.0
Case 1	2.6	0.5	5.0	3.0
Case 2	1.7	0.9	0.5	3.0
Total behavior skills	8.0	1.5	5.0	10.5
Overall score	19.5	2.8	14.1	25.0

SD: standard deviation.

coefficient (ICC). Similarly, the reliability within the medical items and behavioral items were also studied. Proportion test and chi-square tests were done as appropriate. Statistical significance was considered as $p < 0.050$. The selection of the applicants was decided per ranking of the candidates, and the final decision was submitted to admission and registration section. Data was analyzed using SPSS Statistics (IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp).

RESULTS

Fifteen residents participated in the MMI. Ten (66.7%) were females and five (33.3%) males. The mean \pm SD of different domains scores are presented in Table 1.

The mean age of the residents was 25.8 ± 2.2 years. Of the 15 residents, five were recommended for selection and five were kept on the waiting list

as a result of the MMI process. Five were rejected. All residents evaluated the MMI program. The mean international format score was 457.2 ± 50.6 with a minimum score of 402.0 and a maximum of 577.0. The mean medical knowledge score was 4.3 ± 1.2 (range = 3.0–6.0). The mean score of case study one was 2.6 ± 0.5 and 1.7 ± 0.9 for case study two. The mean behavior skills score was 8.0 ± 1.5 . The mean overall MMI score was 19.5 ± 2.8 with a minimum score of 14.1 and a maximum of 25.0.

The reliability (i.e., the ICC) of the scores obtained from the seven items of the MMI was 0.36 (95% confidence interval (CI): -0.31–0.75; $p = 0.110$). However, the ICC of the medical interview was 0.54 (95% CI: 0.45–0.84; $p = 0.090$). This suggested that there was nearly a significant correlation between the items in the medical interview. Similarly, the ICC of the behavioral interview items was 0.18 (95% CI: 0.72–0.69; $p = 0.280$).

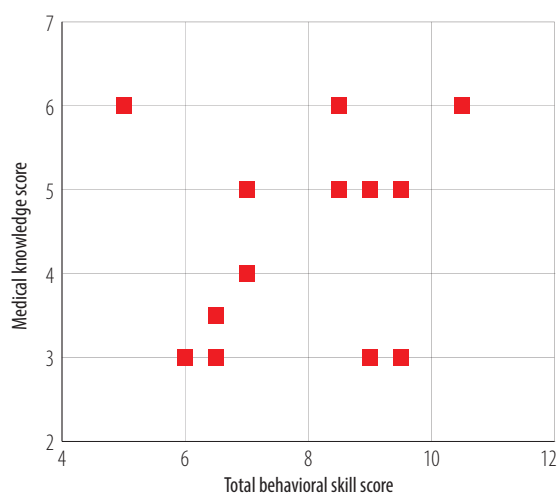


Figure 1: Scatter plot of medical knowledge score with total behavior skills score.

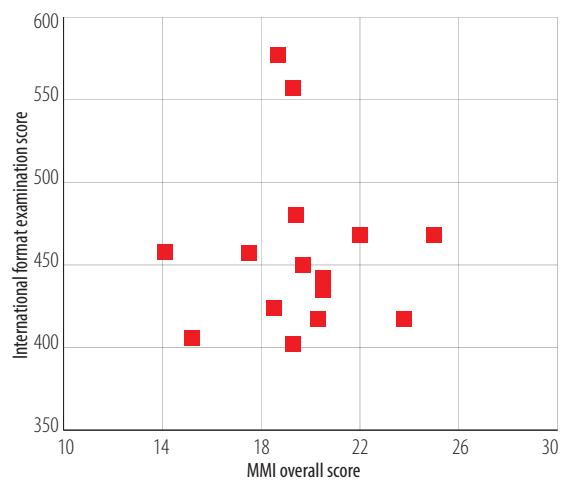


Figure 2: Scatter plot of international format examination score with overall multiple mini-interview (MMI) score.

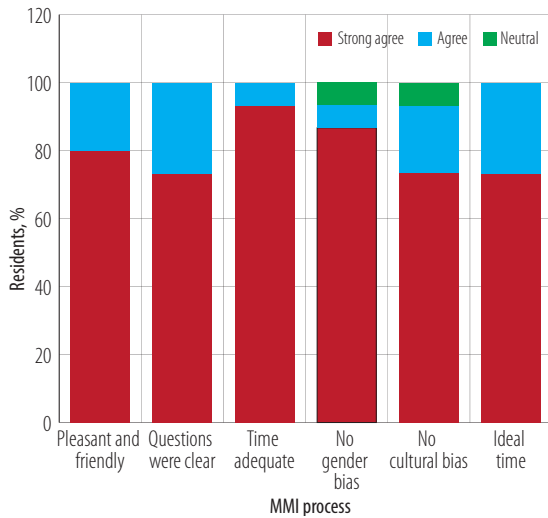


Figure 3: Bar chart showing the residents evaluation of the multiple mini-interview (MMI) process.

The correlation between total behavior items score and the MMI score was $r = 0.135$ ($p = 0.150$). This suggested the MMI score was independent of the behavior score [Figure 1].

Similarly, the correlation coefficient of the international format score and the MMI score was 0.008 ($p = 0.979$), which also suggested the MMI score was independent of the international format score [Figure 2].

All the residents evaluated the MMI process. The results of the evaluation for various questions in the acceptability survey are presented in Figure 3. This figure presents the percentage of residents who have agreed or strongly agreed to these questions. Around 80% of residents agreed or strongly agreed that the MMI process was pleasant and friendly, and 73.3% agreed or strongly agreed the questions were

clear. Regarding the question on whether the time to conduct the MMI was adequate, 93.3% strongly agreed, and 86.7% and 73.3% strongly agreed that there was no gender and cultural bias in the process, respectively.

The residents also compared the MMI process with the traditional process on various items [Table 2]. Regarding favoring the traditional method of screening, 16.7% agreed whereas 100% strongly agreed to favor the MMI screening ($p < 0.001$). Similarly, 16.7% agreed that the traditional method was an enjoyable process while 85.8% agreed or strongly agreed that the process was enjoyable for the MMI method ($p < 0.001$). All residents (100%) agreed or strongly agreed that the MMI was an effective system compared to the traditional method ($p < 0.001$). All residents preferred the MMI method (100%) to the traditional method (0.0%).

DISCUSSION

Our study estimated the reliability of a seven-station MMI model for ORLHNS residency program recruitment in Oman. The generalizability data showed that even with only seven stations, the reliability of our process was about 0.4. However, the consistency coefficient (reliability) of medical stations was about 0.54 (95% CI: 0.45–0.84; $p = 0.090$), suggesting that there was a good level of consistency. Studies have reported higher consistency coefficients due to higher number of stations and over 100 residents. Previous studies have estimated reliability coefficients of 0.73, 0.76, and 0.85 using eight, nine, and 12 stations, respectively.^{1,6,10} Two separate studies

Table 2: Distribution of responses to the evaluation of traditional and multiple mini-interview (MMI) methods by the residents.

Questions	Strongly disagree, %	Disagree, %	Neutral, %	Agree, %	Strongly agree, %	p-value
Was this screening process preferable?						
Traditional method	33.3	33.3	13.7	16.7	0.0	< 0.001
MMI	0.0	0.0	0.0	0.0	100	
Was the process enjoyable?						
Traditional method	16.7	33.3	33.3	16.7	0.0	< 0.001
MMI	0.0	0.0	14.3	42.9	42.9	
Was the process effective?						
Traditional method	0.0	42.9	57.1	0.0	0.0	< 0.001
MMI	0.0	0.0	0.0	42.9	57.1	
MMI vs. traditional	0.0	0.0	0.0	28.6	71.4	

presented a reliability coefficient of 0.7 on an eight-station MMI study (and a consistency coefficient of 0.7),⁸ and 0.67 with the use of 12 stations (and a consistency coefficient of 0.67).¹²

Our study dealt with only 15 residents in seven stations, and the *p*-value was 0.090. Moreover, the ICC = 0.54 was covered within the 95% CI of the above studies. This implied that the reliability was as good as the other studies. The moderate reliability (ICC = 0.54) implied that the seven stations had different concepts for interview and therefore no question duplication. However, some studies have shown ICCs over 0.8, which imply that there may be duplication in different stations.^{6,10} There was no guideline to suggest the minimum number of stations needed to study to assess the reliability and the studies reported so far had a varied number of stations. Hence, the inclusion of seven stations with 15 residents in each station was an arbitrary decision. We hypothesized that the medical domain scores should be independent of the behavior domain score, and our results showed no correlation between them. However, within the behavior stations, the reliability between specialty related grade and the cumulative average grade was 0.71 (*p* < 0.010).

This study has shown that both the interviewers and the interviewees who were previously exposed to standard interviews found this method acceptable. The candidates found the MMI method less taxing and expressed that this method was fairer and more useful in evaluating the non-intellectual traits. Similarly, the examiners also reflected the same viewpoints and endorsed the MMI method instead of the standard interviews. Similar results in accepting the MMI process were expressed in another study based on a group of 74 international medical graduates who applied for a family medicine residency program in Canada.⁵

Based on 484 Canadian and international medical graduates from three residency program in Canada, 88% of the candidates conveyed that they could accurately depict themselves in the MMI process.¹³ Similarly, 74% of the interviewers expressed the view that the MMI process outperformed the traditional method. Moreover, the time spent in the MMI method is as good as the standard interview method. That is, the MMI method needed about 90 minutes per interviewer while this was 80 minutes in the traditional interview method. Thus, the MMI method has been recommended as the better of the two methods.¹³

The main concern of the study was limited candidates in only one residency program. This small number can be attributed to the fact that this study was not designed to hypothesize the applicability of the MMI model. This paper intends to promote the idea of MMI model in countries like Oman, which had been extensively used only in industrialized countries. However, the study recommendations can be strengthened further by involving more centers and more candidates with varying sizes in many centers and different residency programs.

CONCLUSION

The MMI method was found to be acceptable and endorsed as a better method compared to standard interviews by both the candidates and the interviewers. This method can provide reliable results based on seven stations for any residency program.

Disclosure

The authors declared no conflicts of interest. No funding was received for this study. This study was done with the approval of the Otorhinolaryngology Education Committee, Oman Medical Specialty Board. The results of this study were presented as a poster in AMEE 2016, Barcelona, Spain.

Acknowledgements

We are grateful to Oman Medical Specialty Board for the opportunity to conduct this study.

REFERENCES

1. Eva KW, Rosenfeld J, Reiter HI, Norman GR. An admissions OSCE: the multiple mini-interview. *Med Educ* 2004 Mar;38(3):314-326.
2. Batalden P, Leach D, Swing S, Dreyfus H, Dreyfus S. General competencies and accreditation in graduate medical education. *Health Aff (Millwood)* 2002 Sep-Oct;21(5):103-111.
3. Harris S, Owen C. Discerning quality: using the multiple mini-interview in student selection for the Australian National University Medical School. *Med Educ* 2007 Mar;41(3):234-241.
4. Humphrey S, Dowson S, Wall D, Diwakar V, Goodyear HM. Multiple mini-interviews: opinions of candidates and interviewers. *Med Educ* 2008 Feb;42(2):207-213.
5. Hofmeister M, Lockyer J, Crutcher R. The acceptability of the multiple mini interview for resident selection. *Fam Med* 2008 Nov-Dec;40(10):734-740.
6. Eva KW, Reiter HI, Rosenfeld J, Norman GR. The relationship between interviewers' characteristics and ratings assigned during a multiple mini-interview. *Acad Med* 2004 Jun;79(6):602-609.
7. Lemay JF, Lockyer JM, Collin VT, Brownell AK. Assessment of non-cognitive traits through the admissions multiple mini-interview. *Med Educ* 2007 Jun;41(6):573-579.
8. Roberts C, Walton M, Rothnie I, Crossley J, Lyon P, Kumar K, et al. Factors affecting the utility of the multiple mini-interview in selecting candidates for graduate-entry medical school. *Med Educ* 2008 Apr;42(4):396-404.

9. Razack S, Faremo S, Drolet F, Snell L, Wiseman J, Pickering J. Multiple mini-interviews versus traditional interviews: stakeholder acceptability comparison. *Med Educ* 2009 Oct;43(10):993-1000.
10. 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 Aug;43(8):767-775.
11. Reiter HI, Eva KW, Rosenfeld J, Norman GR. Multiple mini-interviews predict clerkship and licensing examination performance. *Med Educ* 2007 Apr;41(4):378-384.
12. Hofmeister M, Lockyer J, Crutcher R. The multiple mini-interview for selection of international medical graduates into family medicine residency education. *Med Educ* 2009 Jun;43(6):573-579.
13. Dore KL, Kreuger S, Ladhani M, Rolfson D, Kurtz D, Kulasegaram K, et al. The reliability and acceptability of the multiple mini-interview as a selection instrument for postgraduate admissions. *Acad Med* 2010 Oct;85(10 Suppl):S60-S63.
14. Harasym PH, Woloschuk W, Mandin H, Brundin-Mather R. Reliability and validity of interviewers' judgments of medical school candidates. *Acad Med* 1996 Jan;71(1 Suppl):S40-S42.
15. Edwards JC, Johnson EK, Molidor JB. The interview in the admission process. *Acad Med* 1990 Mar;65(3):167-177.
16. Elam CL, Andrykowski MA. Admission interview ratings: relationship to applicant academic and demographic variables and interviewer characteristics. *Acad Med* 1991 Sep;66(9 Suppl):S13-S15.
17. Elam CL, Johnson MM. An analysis of admission committee voting patterns. *Acad Med* 1997 Oct;72(10 Suppl 1):S72-S75.
18. Quintero AJ, Segal LS, King TS, Black KP. The personal interview: assessing the potential for personality similarity to bias the selection of orthopaedic residents. *Acad Med* 2009 Oct;84(10):1364-1372.
19. Mann WC. Interviewer scoring differences in student selection interviews. *Am J Occup Ther* 1979 Apr;33(4):235-239.
20. Eva KW. On the generality of specificity. *Med Educ* 2003 Jul;37(7):587-588.
21. Kreiter CD, Yin P, Solow C, Brennan RL. Investigating the reliability of the medical school admissions interview. *Adv Health Sci Educ Theory Pract* 2004;9(2):147-159.
22. Norman G. The morality of medical school admissions. *Adv Health Sci Educ Theory Pract* 2004;9(2):79-82.