

# Multiple Mini Interviews for Selection of Residents and Fellows in a Cardiovascular Center. A 5-year Experience

## *Mini entrevistas múltiples para la selección de residentes y fellows en un centro cardiovascular. Experiencia a 5 años*

MARTÍN A. FASAN<sup>1</sup>, LUCRECIA M. BURGOS<sup>1, MTSAC</sup>, ALAN R. SIGAL<sup>1</sup>, JUAN P. COSTABEL<sup>1, MTSAC</sup>, ALBERTO E. ALVES DE LIMA<sup>1,2, MTSAC</sup>

### ABSTRACT

**Background:** Multiple mini-interviews (MMIs) serve as a model to evaluate non-cognitive skills in the admission process of health care professionals.

**Objective:** The aim of this study was to evaluate the feasibility, reliability and acceptability of the MMI model for the selection of residents and fellows in a cardiovascular center in the past 5 years.

**Methods:** We conducted an observational study including applicants to the cardiology residency program and to the fellowship in Nuclear Medicine and Cardiovascular Ultrasound in 2018, 2019 and 2022. Ten stations were developed to evaluate different non-cognitive domains. Reliability was assessed using G-coefficient. Applicants and interviewers were also surveyed to assess the acceptability of the MMI model and its feasibility in terms of the time required for the process.

**Results:** A total of 75 applicants participated in the MMIs. The G study showed reliability coefficients of 0.62 and 0.61 according to the design. Implementation was feasible; 92% of applicants gave positive reviews to the MMI model, and 90% of interviewers reported they had sufficient time to assess the participants and that the process was not an excessively exhausting.

**Conclusion:** MMIs are a novel method in our setting, demonstrating reliability and a high level of acceptability for evaluating non-cognitive skills in the selection process of applicants to the cardiology residency program and fellowships in a cardiovascular center.

**Keywords:** Cardiology - Internship and Residency - Medical Education

### RESUMEN

**Introducción:** Las mini entrevistas múltiples (MME) son un modelo para evaluar las habilidades no cognitivas en la selección de profesionales ingresantes a instituciones médicas.

**Objetivo:** El objetivo de este trabajo fue evaluar la factibilidad, confiabilidad y la aceptabilidad de las MME para la selección de residentes y fellows en un centro cardiovascular en los últimos 5 años.

**Material y métodos:** Se realizó un estudio observacional, en el cual se incluyeron consecutivamente postulantes a la residencia de Cardiología y a las especialidades de Medicina Nuclear y Ultrasonido en los años 2018, 2019 y 2022. Se desarrollaron diez estaciones para evaluar diferentes dominios no cognitivos. La confiabilidad se evaluó mediante el coeficiente G de generalización. Además, se encuestó a postulantes y entrevistadores para evaluar la aceptabilidad de las MME, y se evaluó la factibilidad en términos de tiempo dedicado al proceso.

**Resultados:** Un total de 75 postulantes participaron de las MME. A partir del estudio G se obtuvieron coeficientes de confiabilidad de 0,62 y 0,61 acorde al diseño. Fue factible su implementación y el 92% de los postulantes valoró de manera muy positiva a las MME. El 90% de los entrevistadores refirió tener suficiente tiempo para evaluar a los participantes y que el proceso no era excesivamente agotador.


**Conclusión:** Las MME son un método novedoso en nuestro medio. Demostraron ser confiables y con un elevado nivel de aceptabilidad para la evaluación de habilidades no cognitivas en el proceso de selección de postulantes a residencia de Cardiología y de subespecialidades en un centro cardiovascular.

**Palabras clave:** Cardiología - Internado y Residencia - Educación médica

REV ARGENT CARDIOL 2023;91:336-342. <http://dx.doi.org/10.7775/rac.v91.i5.20674>

Received: 08/18/2023 – Accepted: 09/29/2023

Address for correspondence: Martín A. Fasan. Instituto Cardiovascular de Buenos Aires, Argentina. Blanco de Encalada 1543. 1428, Buenos Aires, Argentina. E-mail: mfasan@icba.com.ar

 <https://creativecommons.org/licenses/by-nc-sa/4.0/>  
@Revista Argentina de Cardiología

<sup>1</sup> ICBA Instituto Cardiovascular, Buenos Aires, Argentina.

<sup>2</sup> National Scientific and Technological Research Council (CONICET), Buenos Aires, Argentina

## INTRODUCTION

Selecting candidates for medical training programs, as residencies and fellowships, is an essential process that requires institutions to invest significant time and resources. The goal is to choose individuals who possess the cognitive and non-cognitive skills required by each center. (1)

In Argentina, the admission process of residents and fellows typically involves two consecutive stages. The first stage evaluates knowledge through a multiple-choice exam, and the second stage assesses non-cognitive competencies such as professionalism, teamwork, and communication through interviews. (2) These characteristics are better predictors of personal, academic, and professional success than cognitive competencies evaluated through standardized tests. (3-5) However, there is evidence indicating that the traditional interview is not a reliable tool for assessing these competencies. (6)

The reliability of traditional interviews can be undermined by several factors, including variability in interviewer proficiency, the questions posed, possible biases such as benevolence or rigidity, and the context specificity of each interview. (6-9) In addition, if interviewers have preexisting knowledge of a candidate's academic background, it may artificially influence their impression of the candidate. (9) Additionally, there exists significant variation among interviewers, which can be attributed to factors such as their level of experience and potential bias related to culture, age, or gender. Additionally, there exists significant variation among interviewers, which can be attributed to factors such as their level of experience and potential bias related to culture, age, or gender. These inconsistencies have caused the process to be seen as a very labor-intensive and elaborate "lottery." (7)

Considering these limitations, the Multiple Mini Interview (MMI) model is a potential solution for reducing bias and enhancing objectivity by increasing the number of interviews and using standardized questions. (9) Multiple mini interviews provide a more comprehensive assessment of the candidates, preserving validity, acceptability, and reliability. (11-14). These interviews are based on the format of Objective Structured Clinical Examinations (OSCE) and consist of a series of consecutive stations where each interviewer evaluates different aspects of the applicant. (15) This structure ensures standardization and consistency, while allowing for customization according to the needs and expectations of each program. (16)

This model has been implemented in several residency programs, demonstrating high level of reliability and validity (2, 9, 17-20). Its transparent and equitable nature is emphasized, and it is perceived as fair and free of socioeconomic, gender, or cultural biases. (20, 21)

In recent years, our institution has integrated MMIs into the selection process for residencies and subspecialties, including cardiology. (22) This study aims to

assess the reliability, feasibility, and acceptability of the 10-station MMI model in the admission process of cardiology residents and fellows in subspecialties.

## METHODS

### Study design

We conducted an observational and prospective study with psychometric analysis in a high-complexity center specialized in cardiovascular disease. The study included applicants for the cardiology residency program and for the fellowship in cardiovascular ultrasound (CVUS) and nuclear medicine (NM) during 2018, 2019 y 2022. During 2020 and 2021, the MMIs were not conducted following the COVID-19 safety standards.

Applicants were selected based on their performance in a multiple-choice exam created by the Teaching and Research Department of our institution. The exam tested their knowledge, reasoning, and application abilities. This exam was prepared following Galofré's criteria (23) and respecting the following number of questions and proportion of contents: of 150 questions for cardiology residency applicants, 70% corresponded to internal medicine, 10% to gynecology and obstetrics, 10% to pediatrics and 10% to surgery. For CVUS and NM fellowship applicants, the exam had 100 questions on clinical cardiology. Based on the obtained scores, 4 to 5 candidates for each vacant position were invited to the next stage of the MMI by year, following the established order of merit. The candidates were contacted via telephone and email and provided their voluntary consent to participate in the evaluation process.

Professionals from different health areas (physicians of different specialties, nurses, nutritionists, and psychologists) participated as interviewers and were trained in the MMI methodology and received instructions on how to score the applicants.

Ten stations were designed with different scenarios. At each station, a real-world problematic situation was presented to evaluate the applicant's attitudes and qualities relevant to the medical training programs and in line with the institution's vision. Each scenario had three modalities: 1) the interviewer passively observed the applicant interact with a third party, whether it be an actor or another applicant (4 stations); 2) the applicant is presented with a situational vignette and asked to take a position with justification (4 stations); 3) during a semi-structured interview, the applicant is asked about his/her motivation for selecting the specialty, the institution, and extracurricular activities that have provided experience in his/her training as a physician (2 stations). During the MMI process, candidates were assigned an order number to avoid disclosing their affiliations. This measure aimed to minimize any potential bias related to academic performance known to the evaluators.

Different non-cognitive domains were evaluated in each station: teamwork, argumentative skills, professionalism, ethical reasoning, motivation, feedback, acceptance of professional limits and communication skills. Additionally, interviewers were instructed to identify specific "red flags" in candidates' performance that signal a lack of professionalism, such as conflicting attitudes that could potentially exclude the applicant from the selection process (e.g., failure to comply with a patient's advance directive). Based on the scores obtained in the MMI process, a new ranking was prepared for each of the specialties, which was considered for the final selection of applicants to the residency and fellowship programs.

At the end of each day, both interviewers and candidates were invited to complete a web-based satisfaction survey to assess the acceptability of the MMIs as part of the selection process.

In addition, we analyzed the amount of time allocated to each MMI day in relation to the number of applicants to evaluate the feasibility of the method.

### Statistical analysis

We conducted a generalizability theory study to examine the reliability of the assessment tool, an extension of classical reliability theory that permits us to examine the sources of error that affect the scores obtained in an assessment. (24)

The analysis of variance components enables the quantification of sources of variation without the need for complex experimental designs. The G coefficient measures how accurately the objects of study have been differentiated, indicating how well the procedure has classified the objects on a scale of measurement. G coefficients  $\geq 0.80$  indicate a satisfactory level of accuracy for the evaluated model, while coefficients between 0.70 and 0.80 indicate moderate level of accuracy.

It is possible to plan different strategies for the number of assessment occasions, assessment formats, and interviewers needed to obtain reliable results with minimal sampling. This is known as a D-study or Optimization.

In this study, the residents (R) were considered as the object of analysis and the stations (S) as facets to assess the reliability of the tool. The G generalizability coefficient was calculated and alternatives were proposed to optimize the design (D-study) The analysis was performed using EduG 6.1-e software package. (25)

### Ethical considerations

The confidentiality and anonymity of the participants in the study were ensured. Participants were orally invited to take part in the survey and informed of its objectives, with an emphasis on their voluntary participation. The study was conducted following the recommendations of the Declaration of Helsinki and was approved by the institutional review board.

## RESULTS

During 2018, 2019 and 2022, 76 applicants for the cardiology residency program ( $n = 49$ ) and CVUS ( $n = 23$ ) and NM ( $n = 4$ ) fellowships were called to participate in MMIs. Of the 76 applicants, 75 participated in the interviews.

The MMIs took place during a single day each year. The candidates rotated through a circuit of 11 stations (10 station for candidate's evaluation and one rest station) with a total duration of 90 minutes. Before beginning the circuit, the candidates were briefed on the MMI methodology and were randomly assigned a starting station. To avoid biases related to applicants' identities, an order number was assigned to each candidate.

Each station took place in a different room. Candidates had 1 minute to read the vignette situation and the assigned instruction and 7 minutes to complete each station. At the end of the time convened, an audible signal indicated that the participants should move on to the next station. The interviewers used an evaluation form for each candidate featuring a visual analog scale to evaluate non-cognitive domains and rate overall performance at each station. The scores obtained were entered into a table where a corresponding relative value was assigned to each domain and to the overall score (see Appendix). In addition, a free-text commentary field was available to justify the score obtained by the applicant and to note any "red flag" issues that could exclude the applicant from the selection process.

Table 1 shows the different G coefficients and the variance associated with each facet of differentiation according to the different generalizability designs. Two facets were considered: the score obtained including the global score per station and the score without including the global score per station. For the MMI model analyzed, a relative G coefficient of 0.62 was obtained with the global score and 0.61 without the global score.

The optimal number of stations was evaluated from the decision study shown in Table 2. In the conducted analysis, the G coefficient increases to 0.72 when the number of stations is increased to 16.

At the end of the day, the applicants and interviewers filled out a satisfaction survey regarding the MMIs (Figures 1 and 2).

Among applicants, 92% rated the overall MMIs with a score of 4 or 5 on a 1 to 5 scale; 88% responded that they had sufficient time to express their ideas and 81% of the candidates considered that they preferred MMI instead of other types of interviews as a selection method. Most of them responded that MMIs are free of cultural and gender biases.

Among interviewers, 90% responded that they had sufficient time to evaluate the candidates and 90% considered that the day of the interview was not excessively exhausting.

The feasibility of the method was evaluated in terms of the time spent per applicant. In each day of the MMIs, 2 circuits of 10 stations were conducted with a break for the interviewers in between each circuit. With 22 applicants over an estimated time of 210 minutes (90 minutes for the first circuit, a 30-minute break for the interviewers, and 90 minutes for the second circuit), each applicant spent 9.5 minutes on the interview, which is less than the traditional interview time of about 15 minutes. The attendance rate of the

**Table 1.** Generalizability analysis of the 10-station MMI model

	Relative G coefficient	Variance % associated with the facet of differentiation
With global score	0.62	11.7
Without global score	0.61	11.6

applicants was 99%(only one absentee was recorded). All the applicants completed 100% of the stations within the stipulated time.

**DISCUSSION**

This paper describes the past 5-year experience implementing the MMI model for selecting residents and fellows in a cardiovascular center.

According to our expectations, the MMIs were easily implemented with adequate level of reliability and high acceptance by applicants and interviewers as a method to assess non-cognitive competencies in physicians applying for residency and fellowship programs.

Regarding the reliability evaluated through the generalizability analysis, the value of the G coefficient was 0.61, similar to those published in the international literature. Eva et al. (2004) reported a reliability coefficient of 0.65 for a 10-station MMI model. (9) Dore et al. used a 7-station model to select applicants for pediatrics, obstetrics-gynecology and internal medicine residency programs, and obtained a reliability coefficient between 0.55 and 0.72. (26) Furthermore, the G coefficients obtained for the analysis with

and without the global score are similar, indicating that this score does not seem to be associated with a significant source of variation.

The D-study shows that, by increasing the number of stations, the reliability of the model improves, obtaining a G coefficient of 0.72 for 16 stations. In line with Eva, Hofmeister and Roberts, the number of stations has a positive impact on the reliability of the model and is the factor with the greatest influence. (9,27,28) For future selection processes, it may be worth considering designing MMIs with more number of stations. However, this may not be feasible as it would require a larger number of personnel and resources.

In addition, we demonstrated the acceptability of the tool among both applicants and interviewers, as previous studies have reported. Candidates reported that they could express their ideas throughout the MMIs and preferred this evaluation method instead of conventional interviews. This point is likely attributed to the presence of multiple observers, which minimizes the occurrence of subjectivities and biases that are usually present in the traditional interview.

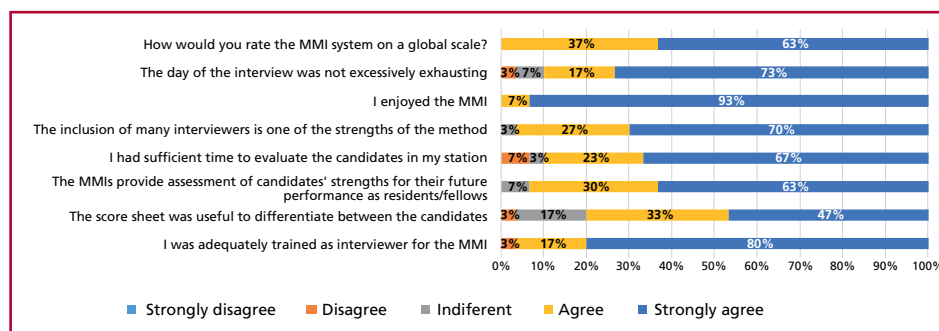
Interviewers found MMIs to be a method that was not excessively exhausting, they had sufficient time to rate the candidates, and most were able to enjoy the experience.

In terms of time spent, MMIs appear to be feasible, with similar or even less time spent on each candidate than the traditional interview, and with a high level of participation. It is also worth noting that MMIs have been successfully performed in our center on 3 different occasions, demonstrating the feasibility of the method.

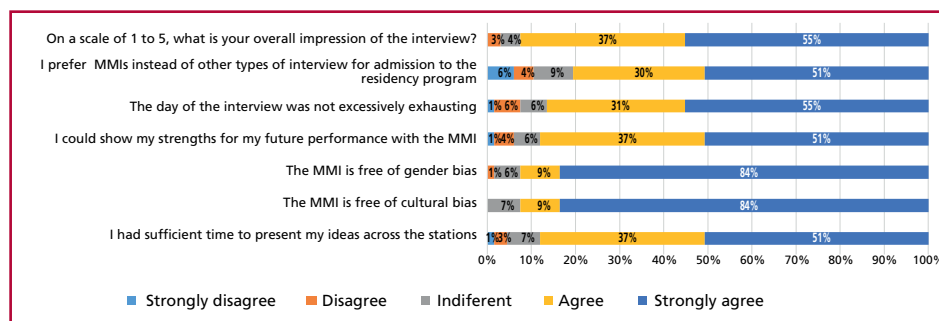
**Table 2.** Decision study (D-study): relative G coefficient for each generalizability design.

Number of stations	R/S'
10	0.62
12	0.66
14	0.69
16	0.72

R': residents; S': stations.



**Fig. 1.** Satisfaction survey with the MMI among evaluators.



**Fig. 2.** Satisfaction survey with the MMI among applicants.



Moreover, we can highlight that this selection method appears to be free of cultural bias.

A limitation of our study is that it was performed in a single private center in Argentina, so the feasibility and acceptability of the method may not be the same for other institutions. In addition, this study includes data from MMIs performed before and after the COVID-19 pandemic, so the context may influence the reliability of the results. It should be noted that the MMI requires an additional effort from health professionals in terms of training in this methodology and in the development of assessment stations and tools. Institutions that adopt this evaluation system must assign a specific human resource to this task.

## CONCLUSION

Our study provides evidence of the feasibility of implementing a 10-station MMI model for the selection of candidates for the cardiology residency program and subspecialties in a cardiovascular center in Argentina. The model showed a high level of acceptance by candidates and interviewers, with an acceptable level of reliability for assessing non-cognitive competencies and could be recommended as a method for selecting professionals.

## Conflicts of interest

None declared.

(See authors' conflict of interests forms on the web).

## REFERENCES

1. Edwards JC, Johnson EK, Molitor JB. The interview in the admission process. *Acad Med.* 1990;65:167-77. <https://doi.org/10.1097/00001888-199003000-00008>
2. Fraga JD, Oluwasanjo A, Wasser T, Donato A, Alweis R. Reliability and acceptability of a five-station multiple mini-interview model for residency program recruitment. *J Community Hosp Intern Med Perspect.* 2013;3:3-4. <https://doi.org/10.3402/jchimp.v3i3-4.21362>
3. Matthews MD, Lerner RM, Annen H. (2019). Non-cognitive amplifiers of human performance: Unpacking the 25/75 rule. In M. D. Matthews and D. M. Schnyer (eds.), *Human Performance Optimization: The science and ethics of enhancing human capabilities*. New York: Oxford University Press.
4. Leininger, Lindsey Jeanne, and Ariel Kalil. "Cognitive and Non-Cognitive Predictors of Success in Adult Education Programs: Evidence from Experimental Data with Low-Income Welfare Recipients." *Journal of Policy Analysis and Management.* 2008;27:521. <https://doi.org/10.1002/pam.20357>
5. Heckman J, Stixrud J, Urzua S. The Effects of Cognitive and Non-cognitive Abilities on Labor Market Outcomes and Social Behavior. *J Labor Ec* 2006;24.
6. Kreiter CD, Yin P, Solow C, Brennan RL. Investigating the reliability of the medical school admissions interview. *Adv Health Sci Educ Theory Pract.* 2004;9:147-59. <https://doi.org/10.1023/B:AHSE.0000027464.22411.Of>
7. Norman G. The morality of medical school admissions. *Adv Health Sci Educ Theory Pract.* 2004;9:79-82. <https://doi.org/10.1023/b:ahse.0000027553.28703.cf>
8. Mann WC. Interviewer scoring differences in student selection interviews. *Am J Occup Ther.* 1979;33:235-9.
9. Eva KW, Rosenfeld J, Reiter HI, Norman GR. An admissions OSCE: the multiple mini-interview. *Med Educ.* 2004;38:314-26. <https://doi.org/10.1046/j.1365-2923.2004.01776.x>
10. Chatterjee A. The interview, revisited. *Clin Teach.* 2018 Feb;15:76-7. <https://doi.org/10.1111/tct.12672>
11. Reiter HI, Eva KW, Rosenfeld J, Norman GR. Multiple mini-interviews predict clerkship and licensing examination performance. *Med Educ.* 2007;41:378-84. <https://doi.org/10.1111/j.1365-2923.2007.02709.x>
12. Roberts C, Walton M, Rothnie I, Crossley J, Lyon P, Kumar K, et al. Factors affecting the utility of the multiple mini-interview in selecting candidates for graduate-entry medical school. *Med Educ.* 2008;42:396-404. <https://doi.org/10.1111/j.1365-2923.2008.03018.x>
13. Lemay JF, Lockyer JM, Collin VT, Brownell AK. Assessment of non-cognitive traits through the admissions multiple mini-interview. *Med Educ.* 2007;41:573-9. <https://doi.org/10.1111/j.1365-2923.2007.02767.x>
14. Kumar K, Roberts C, Rothnie I, du Fresne C, Walton M. Experiences of the multiple mini-interview: a qualitative analysis. *Med Educ.* 2009;43:360-7. <https://doi.org/10.1111/j.1365-2923.2009.03291.x>
15. Kim KJ, Kwon BS. Does the sequence of rotations in Multiple Mini Interview stations influence the candidates' performance? *Med Educ Online.* 2018;23:1485433. <https://doi.org/10.1080/10872981.2018.1485433>
16. Eva KW, Macala C, Fleming B. Twelve tips for constructing a multiple mini-interview. *Med Teach.* 2019;41:510-6. <https://doi.org/10.1080/0142159X.2018.1429586>
17. Farah S, Dewan K, Begum N, Muneer Mian I, Shahzad R. Reliability and acceptability of the multiple mini-interviews for selection of residents in cardiology: Student response. *J Adv Med Educ Prof.* 2021;9:245-6. <https://doi.org/10.30476/JAMP.2020.86698.1260>
18. Yusoff MSB. Multiple Mini Interview as an admission tool in higher education: Insights from a systematic review. *J Taibah Univ Med Sci.* 2019;14:203-40. <https://doi.org/10.1016/j.jtumed.2019.03.006>
19. Al Abri R, Mathew J, Jeyaseelan L. Multiple mini-interview consistency and satisfactoriness for residency program recruitment: Oman evidence. *Oman Med J.* 2019;34(3):218-223. doi:10.5001/omj.2019.42
20. Al Abri R, Mathew J, Jeyaseelan L. Multiple Mini-interview Consistency and Satisfactoriness for Residency Program Recruitment: Oman Evidence. *Oman Med J.* 2019;34:218-23. <https://doi.org/10.5001/omj.2019.42>
21. Finlayson HC, Townson AF. Resident selection for a physical medicine and rehabilitation program: feasibility and reliability of the multiple mini-interview. *Am J Phys Med Rehabil.* 2011;90:330-5. <https://doi.org/10.1097/PHM.0b013e31820f9677>
22. Burgos LM, DE Lima AA, Parodi J, Costabel JP, Ganiele MN, Durante E, et al. Reliability and acceptability of the multiple mini-interview for selection of residents in cardiology. *J Adv Med Educ Prof.* 2020;8:25-31. <https://doi.org/10.30476/jamp.2019.83903.1116>
23. Galofré A, Wright A. Índice de calidad para evaluar preguntas de opción múltiple. *Rev Educ Cienc Salud* 2010;7:141-5.
24. Cardinet J, Johnson S, Pini G. *Applying generalizability theory using EduG*. New York: Routledge; 2009.
25. Swiss Society for Research in Education Working Group. *EDUG user guide*. Neuchâtel, Switzerland: IRDP; 2006.
26. Dore KL, Kreuger S, Ladhani M, Rolfson D, Kurtz D, Kulasegaram K, et al. The reliability and acceptability of the Multiple Mini-Interview as a selection instrument for postgraduate admissions. *Acad Med.* 2010;85:60-3. <https://doi.org/10.1097/ACM.0b013e3181ed442b>
27. Hofmeister M, Lockyer J, Crutcher R. The multiple mini-interview for selection of international medical graduates into family medicine residency education. *Med Educ.* 2009;43:573-9. <https://doi.org/10.1111/j.1365-2923.2009.03380.x>
28. Roberts C, Walton M, Rothnie I, Crossley J, Lyon P, Kumar K, et al. Factors affecting the utility of the multiple mini-interview in selecting candidates for graduate-entry medical school. *Med Educ.* 2008;42:396-404. <https://doi.org/10.1111/j.1365-2923.2008.03018.x>

Appendix

**Table.** Scoring grid by domain and station. The table shows the score assigned to each domain as evaluated at each station.


Station number	Domains							Overall performance	Total score per station
	Motivation toward the specialty	Teamwork	Ethical reasoning	Motivation	Feedback	Communication	Acceptance of professional limits		
1	70					30		50	150
2			50			50		50	150
3		30	30	40				50	150
4		40			40	20		50	150
5			40		20	40		50	150
6		50			30	20		50	150
7		50	30			20		50	150
8			40	30			30	50	150
9		70				30		50	150
10				50		50		50	150
Total by domain	70	240	190	120	90	260	30	50	1500

**Evaluator's form: Station 8**

**Applicant number:**


*On the scales below, mark with a cross your opinion of the candidate's performance in each area.*

**Applicant number:**




Poor Excellent

**Acceptance of professional limits:**




Poor Excellent

**Ethical reasoning:**



Poor Excellent

**Overall performance:**



Poor Excellent

In this space, please explain in your own **words how you constructed the evaluation of the candidate's PERFORMANCE**, trying to identify which were the elements on which you based your score, what you considered relevant, if there was any negative factor that could invalidate the rest, and if you considered any component of non-verbal language.

---

---

**Figure.** Evaluator's form. The figure shows an example of a scorecard, in which interviewers entered their scores on a visual analog scale. This assessment was translated into a numerical value and then entered into the scoring grid. In addition, the interviewers had access to a free-text field to provide further details on the factors considered when assigning a rating and where they could express the presence of certain "red flags". Note that the applicant was identified with a number and not with his or her first and last name to preserve identity and avoid bias.