

Ganiele, María de las Nieves

Múltiples mini entrevistas virtuales como método de selección de residentes

Maestría en Educación para Profesionales de la Salud

Tesis 2023

Cita sugerida: Ganiele M de las N. Múltiples mini entrevistas virtuales como método de selección de residentes [tesis de maestría] [internet]. [Buenos Aires]: Instituto Universitario Hospital Italiano de Buenos Aires; 2023 [citado AAAA MM DD]. 85 p. Disponible en: <https://trovare.hospitalitaliano.org.ar/descargas/tesisytr/20240130114343/tesis-ganiele-maria.pdf>

Este documento integra la colección Tesis y trabajos finales de Trovare Repositorio del Institucional del Instituto Universitario Hospital Italiano de Buenos Aires y del Hospital Italiano de Buenos Aires. Su utilización debe ser acompañada por la cita bibliográfica con reconocimiento de la fuente.

Para más información visite el sitio <http://trovare.hospitalitaliano.org.ar/>



**Múltiples Mini-Entrevistas virtuales
como método de selección de residentes**

por

María de las Nieves Ganiele

Director: Prof. Dr. Carlos Alberto Brailovsky

Co-directora: Dra. María Dolores Arceo

**Presentado en cumplimiento de los requisitos para la obtención del Título de
Magíster en Educación para los Profesionales de la Salud**

ante el

Instituto Universitario del Hospital Italiano de Buenos Aires

Buenos Aires, 17 de abril de 2023

Dedicatoria

A Quimi, a Pablo y al nuevo integrante "en camino", quienes son mi sostén, mi alegría y me acompañan en mi camino de crecimiento profesional.

A mis hermanas y a mis padres, Miguel y Noemí que me han animado a estudiar desde que tengo memoria.

Agradecimientos

Al Prof. Dr. Carlos Alberto Brailovsky: co-director de la Maestría en Educación para Profesionales de la Salud y director de mi tesis, por su amabilidad en compartir conmigo toda su expertise y su paciencia. Guardaré siempre un gran recuerdo del tiempo y trabajo compartido para llevar esta tesis a buen puerto.

A la Dra. Dolores "Loli" Arceo: co-directora de mi tesis, colega, promotora innata de proyectos e ideas superadoras que siempre me ha tenido presente para formar parte en sus equipos de trabajo, pero principalmente, una amiga.

Al Prof Dr. Eduardo "Mono" Durante: director de la Maestría en Educación para Profesionales de la Salud, quien desde mis primeros pasos como residente ha sido de gran ayuda e inspiración en mi camino de formación, alentándome a innovar y a hacer el mejor esfuerzo en la enseñanza de nuestra disciplina.

Al Dr. Fernando R. Vázquez: quien dedicó su tiempo y paciencia en asesorarme en aspectos estadísticos donde él es un gran experto.

A los médicos/as residentes y evaluadores/as: que no solo participaron en esta edición de las MME sino que año a año (desde que empezamos en 2016) colaboran en esta metodología de selección que requiere del trabajo conjunto y buena predisposición de todos para llevarla adelante con éxito.

A quienes conforman el departamento de posgrado del IUHI: que apoyan los proyectos innovadores en educación y siempre han sido muy amables conmigo, para este y otros proyectos.

Introducción¹

Uno de los caminos que habitualmente recorre un médico una vez finalizados sus estudios de grado consiste en ingresar a una residencia para completar su formación en alguna especialización reconocida por su país de origen. En el marco de ese programa educativo, los residentes realizan diversas actividades profesionales de complejidad y responsabilidad crecientes bajo supervisión de tutores docentes en instituciones de servicios de salud. El funcionamiento de los sistemas de salud de muchos países asienta en gran medida en el trabajo de los médicos residentes. Es por ello que los ministerios de salud invierten anualmente una suma de dinero importante para cubrir estos puestos, ya que este es el ámbito del cual surgen las nuevas generaciones de especialistas (Organización Panamericana de la Salud. Área de Sistemas de Salud basados en la Atención Primaria de Salud, 2011). El ingreso a la residencia es un momento clave en la educación médica ya que suele haber competencia por las vacantes, lo que varía según la región y especialidad. Ciertas instituciones de formación (privadas como públicas) son muy requeridas ocupando rápidamente sus cargos y generalmente una vez que se ingresa en un programa la mayoría de los médicos se gradúa finalmente como especialista (Fortin et al., 2016).

Dichos programas de residencia tienen como fin formar médicos competentes que logren una práctica independiente, segura y de gran calidad en su área de especialización (Roberts et al., 2018). La competencia clínica se define como “la capacidad de un profesional de utilizar su buen juicio así como también los conocimientos, habilidades y actitudes asociadas a la profesión para solucionar los problemas complejos que se presentan en el campo de su actividad profesional” (Brailovsky, 2001; Kane, 1992). Como marco de referencia el Royal College of Physicians and Surgeons of Canada ha desarrollado el modelo CanMEDS (Canadian Medical Education for Specialists) y CanMEDS-Medicina de Familia que identifica las habilidades que requieren los médicos para responder a las necesidades de las personas a las que atienden. Estas habilidades están agrupadas bajo siete roles: médico experto, buen comunicador, colaborador capaz de trabajar

¹ Aclaración al lector: Se ha decidido por cuestiones de legibilidad conservar el uso tradicional del genérico neutro y plural en masculino. Es decir que cuando en el texto se utilicen términos o pronombres, por ejemplo, "médicos" nos estaremos refiriendo tanto a médicos como médicas. Siempre que sea posible se emplearán palabras que incluyan a todo el universo de personas.

en equipo, líder y con capacidad de gestión, promotor de la salud, académico a través de la educación continua y por último, profesional, demostrando compromiso con los pacientes, la profesión y la sociedad ejerciendo una práctica ética de la disciplina (Frank et al., 2016). De esta forma existen atributos, relacionados con variables cognitivas y no cognitivas, necesarios para una práctica profesional exitosa. En base a la evidencia actual, obtenida principalmente de países de Europa, Canadá y Estados Unidos, estos atributos pueden ser evaluados con diferentes estrategias ya desde el momento de la selección de estudiantes que ingresan a la universidad (grado) o de postulantes en el ingreso a programas de residencia (posgrado). Es así que el proceso de selección debería apreciar no sólo las calificaciones académicas (atributos cognitivos) sino también realizar la valoración de características y aptitudes personales (atributos no cognitivos) que puedan adecuarse mejor al perfil del estudiante/residente buscado (Albanese et al., 2003; Lemay et al., 2007; Roberts et al., 2018).

El foco de interés de esta tesis se relaciona con el proceso de selección de postulantes para residencias médicas. En Argentina como regla general pocos elementos no cognitivos se incorporan en esta instancia y se suelen utilizar instrumentos de evaluación que miden únicamente atributos cognitivos (promedio de la carrera de grado y puntaje del examen de conocimiento) (*Ingreso a Residencias*, 2020b). En muchos casos las instituciones, en su mayoría hospitales universitarios o privados, suman al proceso una entrevista tradicional de tipo panel como un intento de complementar la evaluación de conocimientos y con el objetivo de “conocer” aspectos no académicos de los postulantes. Hace tiempo que la evidencia sugiere que la entrevista tradicional (estructurada o semiestructurada) tiene grandes limitaciones en relación a los sesgos de los evaluadores y presenta confiabilidad y validez cuestionables para la instancia de selección de residentes (Patterson et al., 2016; Sklar et al., 2015). A pesar de que ha sido demostrado que los logros académicos proporcionan un buen predictor del desempeño a lo largo de la formación no resulta tan clara la relación entre el éxito durante la educación y el posterior durante la práctica clínica. No puede asumirse que quienes tengan altas calificaciones se convertirán en médicos competentes al atravesar una residencia, algunos trabajos concluyen que sería necesario presentar ciertos atributos de tipo

no cognitivos desde un principio (Hamdy et al., 2006; Kulatunga-Moruzi & Norman, 2002; Patterson & Ferguson, 2010).

Es así que las Mini Múltiples Entrevistas (MME) se desarrollaron con el objetivo de compensar la falta de validez psicométrica de las entrevistas tradicionales en la evaluación de atributos no cognitivos (Eva et al., 2004). En su inicio fueron aplicadas en el ingreso a la carrera de medicina y desde entonces su empleo se expandió, utilizándose en otras carreras como enfermería, odontología, psicología, veterinaria y también en la selección de residentes, con buenos resultados de confiabilidad y aceptabilidad (Yusoff, 2019). Desde 2016, en base a la creciente evidencia en relación a la confiabilidad de las MME y con la aprobación del Área de Residencias y Becas del Departamento de Posgrado del Instituto Universitario del Hospital Italiano de Buenos Aires (IUHIBA), se decidió reemplazar la entrevista semiestructurada tradicional por este método para la evaluación de atributos no cognitivos en la selección de residentes de Medicina Familiar (Durante et al., 2017). Desde entonces se han sumado otras residencias. En 2020 por las necesidades de distanciamiento físico de la pandemia por COVID-19 se realizaron de manera virtual y se mantienen en esta modalidad hasta el presente por su factibilidad (Carroll Turpin et al., 2021).

En la actualidad está bien establecido que las MME son útiles en la selección de ingresantes y se dispone de evidencia fuerte relacionada a su confiabilidad, validez y aceptabilidad (Roberts et al., 2018; Yusoff, 2019). Sin embargo, la literatura ofrece evidencias no concluyentes acerca de cuáles son los constructos que se miden en las MME (Reiter & Eva, 2018). En el ámbito de la educación y la psicometría se suele utilizar el término constructo para hacer alusión a términos creados para hacer referencia a fenómenos complejos que consisten en una constelación de características o conductas que no se pueden observar de manera directa (Crowl, 1996). Ejemplos de constructos son creatividad, desarrollo humano, aptitud académica. La intención de las MME es medir algún constructo relacionado con la competencia profesional (Reiter & Eva, 2018), hay publicaciones que a través del análisis factorial sugieren que los puntajes obtenidos en las MME colapsan en dos factores, uno relacionado con habilidades comunicacionales y otro más difuso con la competencia personal (Oliver et al., 2014).

Esta tesis tiene como objetivo principal evaluar si la medición de las MME corresponde a un constructo unifactorial o multifactorial con capacidad de diferenciar los atributos no cognitivos estipulados en la tabla de especificaciones al inicio de su planificación. De manera secundaria analizar qué valoran los evaluadores al momento de puntuar, la confiabilidad, validez y aceptabilidad de las mismas en la selección de residentes en un hospital de gestión privada de la Ciudad Autónoma de Buenos Aires.

Resumen

Introducción: Las Múltiples Mini Entrevistas (MME) son un método para valorar aspectos no cognitivos claves en el proceso de selección de residentes.

Objetivos: Evaluar si las MME miden un constructo multi o unifactorial. Explorar percepciones de los evaluadores frente a las actitudes de los postulantes. Establecer su confiabilidad, validez y aceptabilidad para la selección de residentes en el Hospital Italiano de Buenos Aires.

Metodología: Estudio descriptivo, analítico, estrategia cuanti-cualitativa. A partir de los datos de las MME virtuales de 2022 se realizó un análisis factorial exploratorio (AFE) y análisis temático del contenido (comentarios escritos y entrevistas semiestructuradas a evaluadores). Se exploró la confiabilidad (coeficiente alfa de Cronbach) y generalizabilidad (coeficiente G), validez y la aceptabilidad con un cuestionario anónimo y voluntario.

Resultados: Participaron las residencias de cirugía general, medicina familiar y tocoginecología (38 evaluadores y 76 postulantes). El AFE obtuvo un constructo unifactorial y del análisis cualitativo de la información contextual de las puntuaciones surgieron 4 temas principales (biografía, personalidad, comunicación y ejecución) que confluyeron en dos dimensiones, la primera impresión y resolución. Tanto el coeficiente alfa de Cronbach como el G fueron de 0,72. Se obtuvo evidencia de diversas fuentes que apoyó la validez. El cuestionario fue respondido por el 84,2% de los postulantes y el 60,5% de los evaluadores, con una valoración positiva del método.

Conclusión: Las MMEv miden un constructo unidimensional. Los evaluadores valoran principalmente la primera impresión y la resolución de la consigna. Resultó un método confiable, válido y bien aceptado para la selección de residentes.

Palabras clave: Internado y residencia, Educación Médica, Educación de

Tabla de contenidos

Dedicatoria y agradecimientos	i
Introducción	ii
Resumen	vi
Tabla de contenidos	vii
TEXTO DE LA TESIS	
Fundamentación	1
Proceso de selección de residentes en Argentina	1
Evaluación de los atributos no cognitivos a través de entrevistas	2
Múltiples mini-entrevistas: características y utilidad	4
Experiencia local en la implementación y adaptación a la virtualidad	8
Planteo del problema y justificación de la investigación	9
Objetivos de la investigación	10
Metodología	11
Tipo de estudio, diseño y recolección de datos	11
Población y ámbito del estudio	11
Desarrollo e implementación de las Múltiples Mini Entrevistas virtuales	12
Análisis	14
Objetivo primario: Análisis factorial exploratorio	14

Objetivos secundarios	15
Análisis de contenido	15
Confiabilidad	16
Validez	17
Aceptabilidad	18
Consideraciones éticas	18
Resultados	18
Objetivo primario: Análisis factorial exploratorio	21
Objetivos secundarios	23
Análisis de contenido	23
Definición de códigos	24
Dimensiones analíticas	28
Confiabilidad	31
Validez	32
Aceptabilidad	33
Discusión	38
Conclusión	43
Anexo I Tabla especificaciones	i
Anexo II Ejemplos de estaciones y ficha de evaluación	ii
Anexo III Guía de preguntas para entrevista a evaluadores	xi
Anexo IV Cuestionario de aceptabilidad	xiii
Anexo V Consentimiento informado oral	xviii

Fundamentación

Proceso de selección de residentes en Argentina

En Argentina el título de médico otorgado por las universidades habilita para el ejercicio de la profesión, pero para ser especialista es necesaria la certificación. Hay diferentes opciones para obtenerla, la mejor es la formación a través de un programa de residencia. Con el paso del tiempo el balance entre la oferta y la demanda de puestos para ingresar a una residencia se invirtió. Hace 25 años la oferta sólo alcanzaba para el 30% de los médicos graduados, en el 2015 este número se equiparaba (Duré, 2015) y en los últimos años la cantidad de cargos ofrecidos anualmente supera al total de graduados. Se ha observado una tendencia de cobertura de cargos total que no supera los $\frac{2}{3}$ de la oferta a nivel país (con variaciones en ciertas especialidades y regiones) y la presencia de profesionales extranjeros entre quienes intentan acceder a la formación llega al 25% (Diéguez, 2020; Ministerio de Salud de la Nación y Desarrollo Social. Dirección Nacional de Capital Humano, 2019). Este panorama varía según la especialidad, algunas de las más requeridas logran una cobertura del 100% a lo largo de los años como anestesiología, diagnóstico por imágenes, oftalmología, neurocirugía y otorrinolaringología. Otras como cardiología, ortopedia y traumatología, clínica médica y pediatría cuentan con más de un postulante que rinde por cargo y logran alcanzar una cobertura superior a la media nacional aunque no completa. En otras consideradas prioritarias (medicina familiar y general, neonatología, terapia intensiva infantil y de adultos) el 60% de los cargos permanecen vacantes porque los graduados buscan otras instancias de formación o directamente otras especialidades. A su vez hay diferencias en la ocupación de cargos a lo largo del país según la región o localidad y en más detalle en cada área hay instituciones (tanto públicas como privadas) más requeridas que otras (Ministerio de Salud de la Nación y Desarrollo Social. Dirección Nacional de Capital Humano, 2019).

En el año 2011, el Ministerio de Salud de la Nación implementó el examen único (EU) de selección en provincias de la región patagónica y con el paso de los años fue involucrando a un mayor número de instituciones estatales y del subsector privado de casi todo el país (a excepción de la provincia de Córdoba). El EU nació a

partir del Consejo Federal de Salud como una iniciativa que tiene como visión una distribución más equitativa de la oferta de formación de especialistas (Duré, 2015). Sin embargo, este examen es solo una herramienta de evaluación que no busca determinar niveles de competencia clínica sino un orden de mérito y luego cada provincia o institución participante conforma su concurso (provincial, regional o institucional) con variaciones en los requisitos e instancias del sistema de selección (*Ingreso a Residencias*, 2020a). De esta forma el IUHIBA se sumó al EU federal en el año 2019 manteniendo su propio concurso con una instancia posterior de entrevista personal.

El EU es un examen con preguntas de opción múltiple que permite evaluar el conocimiento de tipo sabe y sabe cómo (de acuerdo a la pirámide de Miller) (Miller, 1990) y establecer un ranking. En esta instancia y con las limitaciones propias del instrumento se acepta que tenga una adecuada confiabilidad que ha sido probada en publicaciones previas (Ministerio de Salud de la Nación y Desarrollo Social. Dirección Nacional de Capital Humano, 2019; Serralunga et al., 2017). Frente a esta situación donde un examen evalúa sólo los atributos cognitivos varias instituciones implementan una instancia posterior de entrevista para hacer una apreciación de los atributos no cognitivos de sus posibles ingresantes. Teniendo presente que el fin último es formar profesionales competentes, ya desde el momento de la selección vale la pena valorar no sólo las calificaciones académicas sino también atributos no cognitivos, esas aptitudes personales relacionadas con el futuro ejercicio. El punto clave es cuál es la mejor forma de seleccionar a los futuros residentes dentro de un grupo de postulantes altamente calificados.

Evaluación de los atributos no cognitivos a través de entrevistas

Es sabido que hay competencias específicas de cada especialidad médica y competencias transversales que son relevantes en el ejercicio de la profesión en general (Frank et al., 2016). La selección de qué atributos no cognitivos evaluar debe adecuarse al perfil del residente que la institución establezca. Los más mencionados en la bibliografía son: trabajo en equipo, responsabilidad, ética, habilidades sociales, comunicación, pensamiento crítico y profesionalismo.

En los últimos 50 años la psicometría fue ganando relevancia y la tendencia alrededor del mundo para la valoración de los atributos no cognitivos fue cambiando progresivamente desde la utilización de métodos de evaluación menos estructurados (entrevistas no estructuradas, cartas de referencia y declaraciones personales) hacia modelos con mayor respaldo de la evidencia. De esta manera la selección busca basarse en procedimientos equitativos, surgidos de la validez y confiabilidad de los instrumentos utilizados y que se adecue a la capacidad de cada institución de ofrecer y brindar una enseñanza de calidad para el desarrollo de la salud de la sociedad (Brailovsky & Centeno, 2012).

La entrevista semiestructurada con un panel de entrevistadores es ampliamente utilizada, sin embargo, este formato de entrevista no posee las mejores características psicométricas. Existe una gran variabilidad en cómo es llevada a cabo, carece de confiabilidad y es un mal predictor de lo que va a pasar a futuro con quienes son aceptados en los programas (Edwards et al., 1990; Kreiter et al., 2004; Siu & Reiter, 2009). Otro punto cuestionable es la especificidad de contexto: el desempeño de un individuo es altamente variable según el contexto, como se desenvuelva en una situación no predice cómo lo hará en otras (Eva, 2003). Patterson et al. (2016) en una revisión sistemática sobre métodos de selección en educación médica de grado y posgrado analiza publicaciones desde 1997 a 2015 evaluando la confiabilidad, validez, aceptabilidad, aspectos procedimentales y costo-efectividad de los mismos. Esta revisión sugiere que hay métodos de selección más efectivos y justos tales como el desempeño académico previo, las múltiples mini-entrevistas, pruebas de juicio situacional y centros de selección en comparación a la entrevista tradicional, referencias y declaraciones personales (Patterson et al., 2016).

Con el fin de mejorar estas características psicométricas en la Universidad McMaster (Ontario, Canadá) se desarrollaron las Múltiples Mini Entrevistas (MME) para el ingreso a la carrera de medicina (Eva et al., 2004). A diferencia de las entrevistas tradicionales, donde los postulantes interactúan con un panel de entrevistadores (compuesto por 3 personas habitualmente) por 20-30 minutos, en las MME mayor número de entrevistadores pueden interactuar individualmente con los postulantes en un tiempo más acotado (8-10 minutos). Como resultado, son

capaces de desarrollar un sentido más amplio de la distribución de respuestas y la evaluación por parte de cada uno es independiente de la de los demás. En la actualidad son utilizadas por todas las facultades de medicina de ese país y adoptadas ampliamente alrededor del mundo (Inglaterra, Estados Unidos, Australia, países de Europa y Medio Oriente) en el proceso de selección de aspirantes a diferentes carreras de la salud (medicina, enfermería, odontología, farmacia, etc.) y a residencias médicas (Rees et al., 2016; Yusoff, 2019). A nivel del posgrado son utilizadas en Inglaterra, Canadá, Australia, Estados Unidos y se realizaron trabajos piloto en Japón, Pakistán y Medio Oriente. En Latinoamérica se encuentra publicado un trabajo realizado en la admisión a la carrera de Medicina en una Universidad de San Pablo, Brasil como primera experiencia en ese país que obtuvo resultados favorables (Daniel-Filho et al., 2017) y nuestra experiencia en posgrado para la selección de residentes de cardiología en el Instituto Cardiovascular de Buenos Aires (ICBA) (Burgos et al., 2020). Al momento las MME solo se utilizan en el ámbito de la educación en profesiones de la salud.

Múltiples mini-entrevistas: características y utilidad

Las MME fueron desarrolladas en base al modelo del examen clínico objetivo estructurado (ECO) ya que utiliza varias estaciones de evaluación independientes típicamente en un circuito reglado por tiempo, pero en este caso no son objetivas ni abordan temas clínicos (Eva et al., 2004). En las MME, los postulantes atraviesan un circuito de estaciones encontrándose en cada una de ellas una entrevista breve donde deben resolver situaciones problemáticas simuladas con el fin de evaluar los atributos no cognitivos establecidos en el diseño inicial. Considerando que un postulante puede tener diferente rendimiento en cada situación, el uso de múltiples escenarios resulta más consistente desde el punto de vista psicométrico teniendo en cuenta las bases de la teoría educativa y de la evaluación. Es decir, un postulante puede tener bajo rendimiento en una estación pero mejorar sustantivamente en otra, ya que personas con diferente camino recorrido pueden demostrar sus cualidades en diferentes escenarios y destacarse más en unos que en otros. Adicionalmente se puede asegurar que los puntajes son totalmente independientes porque cada entrevistador permanece en un cuarto separado del resto mientras está en funcionamiento el circuito de entrevistas. En definitiva, la principal ventaja de las

MME es que las múltiples entrevistas (estaciones) diluyen el efecto del azar, los sesgos situacionales y de los entrevistadores (Eva et al., 2004).

Se utiliza el término entrevista pero en realidad no todas las estaciones son de entrevista, sino que el formato es flexible y se pueden diseñar diferentes modalidades:

- Situación hipotética: En este caso hay interacción entre el entrevistador y el postulante, a quien se le presenta una situación problemática y se solicita que tome posición con fundamentación en base a preguntas orientadoras. Estas situaciones deben ser lo más parecidas posible a la realidad y pueden abarcar temáticas muy variadas de la vida diaria, relacionadas o no con la profesión (Ej. actualidad, arte, planificación de vacaciones, discrepancia con un compañero de trabajo, etc.).
- Escenarios basados en la experiencia: En este caso hay interacción entre el entrevistador y el postulante, a quien se pide que relate a alguna situación que le haya ocurrido y que argumente en relación a su actuación, aprendizaje, etc.
- Interacción con un tercero: el entrevistador actúa como observador de la interacción del postulante con un actor/actriz (comparable al paciente simulado del ECOE) o con otro postulante.
- Interacción con varios postulantes: el entrevistador actúa como observador de la interacción de un grupo pequeño de postulantes (4-5 personas) que deben resolver una tarea de manera conjunta.
- Entrevista semi estructurada: entrevista de tipo tradicional en donde el entrevistador posee una guía de preguntas a realizar con el fin de conocer la motivación en la selección de la especialidad, de la institución, experiencia, desafíos personales y actividades extracurriculares que puedan ser relevantes para el puesto.

Esta amplitud de opciones para su desarrollo hace menos probable que los postulantes preparen con anticipación su desempeño, y aunque practiquen alguna respuesta en su mayoría tendrán que actuar espontáneamente frente a las diferentes situaciones que se les presenten. Los evaluadores poseen un formulario

de puntuación estandarizado para calificar a los participantes y así obtener un puntaje por estación y un puntaje por cada atributo evaluado (ej. habilidades de comunicación, pensamiento crítico, trabajo en equipo). A diferencia del ECOE donde los estudiantes deben resolver una situación clínica, acá no existen respuestas correctas o incorrectas para los escenarios propuestos, solo se evalúan las conductas frente a la situación planteada ya que lo que se valora es la capacidad de expresar sus ideas claramente (Eva et al., 2004, 2019). Partiendo de que el conocimiento médico ya fue evaluado previamente con otro tipo de instrumento (en nuestro caso en el Examen Único) en esta instancia no se espera que los postulantes tengan conocimientos específicos tales como diagnosticar o tratar una patología.

Con respecto a los puntos prácticos en el diseño e implementación de las MME, el primer paso es confeccionar la tabla de especificaciones en base a los atributos no cognitivos definidos como relevantes para el perfil de ingresantes de la institución (Eva et al., 2019). Los 10 atributos más evaluados según lo reportado en la bibliografía son profesionalismo, habilidades de comunicación, trabajo en equipo, ética y moral, pensamiento crítico y resolución de problemas, motivación para la carrera, empatía, habilidades de gestión, resiliencia y habilidades interpersonales (Yusoff, 2019). Está en discusión si realmente se puede evaluar cada uno de esos atributos separadamente o si verdaderamente se evalúa un único constructo global mayormente relacionado con lo comunicacional o desempeño integral. Con respecto a este punto la evidencia publicada es heterogénea y será una de las cuestiones de análisis de la presente tesis.

En general, el número de estaciones utilizadas varía entre 7 y 12, es sabido que la confiabilidad aumenta con cada estación adicional pero se observó que la confiabilidad test-retest llega a una meseta con 8-10 estaciones (Eva et al., 2009, 2019). Aumentar la cantidad de estaciones tiene más impacto en la confiabilidad que aumentar la duración de las mismas o la cantidad de evaluadores por estación (Knorr & Hissbach, 2014; Rees et al., 2016). La mayoría de las instituciones adhiere a que lo más apropiado es utilizar solo un entrevistador por estación (Hecker & Violato, 2011; Rees et al., 2016). Es clave realizar una buena preparación de los evaluadores para su rol, en donde reciban instrucciones claras de cuáles son los

atributos no cognitivos que evaluarán, estimular a que usen toda la amplitud de la escala de calificación y realizar una prueba de la estación asignada. Al realizarse múltiples entrevistas la responsabilidad de que una persona ingrese o no a la residencia es repartida entre todos los evaluadores, por lo cual pueden sentirse más libres de asignar una puntuación baja ya que el postulante puede recuperarse si tiene buen desempeño en otras estaciones. También se alienta a pensar sus primeras calificaciones como provisionarias y ajustarlas si al observar el desempeño de los siguientes consideran que fueron demasiado indulgentes o estrictos. Esta calibración suele ocurrir luego de interactuar con 3 a 5 postulantes (Eva et al., 2019).

Al diseñar el contenido de las estaciones se prefiere evitar proponer que se imaginen siendo miembros de la profesión en la cual quieren ingresar, ejemplo "imaginar que es médico especialista en cardiología y está con un paciente con tal diagnóstico". Esto se basa en que aún no están preparados para ese rol y que lo aprenderán en el programa de formación al que van a ingresar. Se sugiere priorizar estaciones que permitan a los postulantes demostrar cualidades de tipo no académicas al reflexionar sobre situaciones de la vida diaria que pudieran haber vivido (ej. siendo estudiantes, amigos, familiares) (Eva et al., 2019).

Frente a la situación de maximizar las cualidades de las MME con los recursos disponibles lo que tiene mayor prioridad de todo lo mencionado hasta el momento es el muestreo independiente, que permite múltiples observaciones individuales de las habilidades de un mismo postulante (Eva et al., 2019).

Debido a la expansión en su utilización, principalmente en el grado, hay publicadas diferentes revisiones sistemáticas que valoran las características psicométricas de las MME. El análisis se realiza siguiendo el concepto de utilidad de un instrumento de evaluación definido por Van der Vleuten a partir de la siguiente ecuación:
$$\text{Utilidad} = (\text{Confiabilidad} \times \text{Validez} \times \text{Impacto Educativo} \times \text{Aceptabilidad}) \times \frac{1}{\text{Costo o factibilidad}}$$
 (Van Der Vleuten, 1996). Dentro de los principales resultados hay evidencia a favor de la confiabilidad, factibilidad, validez y amplia aceptabilidad por parte de postulantes y evaluadores. Teniendo presente cuál es el propósito de las MME el impacto educativo no tiene gran peso al medir su utilidad. La confiabilidad reportada varía entre moderada a alta y la mayoría de las

publicaciones la expresa a través del coeficiente alfa de Cronbach (derivado de la teoría clásica de los tests) o del coeficiente G (de la teoría de la generalizabilidad). De esta manera el coeficiente alfa de Cronbach (consistencia interna de los ítems del instrumento de evaluación) publicado se encuentra entre 0,69 – 0,98 y el coeficiente G (proporción de variabilidad en los puntajes de los individuos, atribuible a sus diferencias sistemáticas en conocimiento, habilidades y experiencias) entre 0.55 - 0.72 (Rees et al., 2016; Roberts et al., 2018; Yusoff, 2019). Es un método con buena aceptación por parte de postulantes y evaluadores, lo consideraron una herramienta de selección justa aunque algunos estudios reportan que puede ser estresante para ciertos aplicantes (Kumar et al., 2009; Rees et al., 2016). Se lo considera libre de prejuicios de género y como un método aceptable en un contexto multicultural (Dore et al., 2010; Hofmeister et al., 2008; Pau et al., 2016; Yusoff, 2019). Al existir evidencia acumulada por la experiencia de diferentes instituciones se puede concluir que las MME pueden ser diseñadas como un método de evaluación válido. A partir de encuestas realizadas a los participantes se obtuvieron resultados muy positivos, más del 90% consideraba que lo abordado en las estaciones era relevante para su futuro entrenamiento (Cameron & Mackeigan, 2012; Dowell et al., 2012; Rees et al., 2016). El consenso internacional sobre admisiones publicado luego de la Conferencia de Ottawa de 2010 concluyó que las MME eran el único método para medir atributos no cognitivos con evidencia de confiabilidad test-retest y validez predictiva (Prideaux et al., 2011). Por otro lado la factibilidad puede llegar a ser un problema en términos de costo efectividad lo que podría restringir su uso pero se requiere mayor evidencia al respecto (Roberts et al., 2018).

Las revisiones sistemáticas realizadas por Yusoff (2019), Rees et al. (2016) y Pau et al. (2013) concuerdan en que las MME resultan ser un instrumento de selección factible, confiable, válido y libre de prejuicio, y que correlaciona mal con atributos no cognitivos como rasgos de personalidad, atributos rurales (Eley et al., 2009) e inteligencia emocional. En su revisión Yusoff (2019) remarca que falta evidencia que evalúe su correlación con atributos no cognitivos importantes como la salud mental, la ética y el profesionalismo.

Experiencia local en la implementación de las MME y adaptación a la virtualidad

En el nivel de posgrado contamos con experiencia local con MME para la selección de residentes en el IUHIBA y en el ICBA. Nuestra experiencia reafirma el uso de las MME, como un modelo confiable para evaluar atributos no cognitivos y con buena aceptabilidad por parte de postulantes y evaluadores (Burgos et al., 2020; Durante et al., 2017).

Como se mencionó previamente, por las limitaciones de la pandemia por COVID-19 en 2020 y 2021 las entrevistas se realizaron de manera virtual. En 2022 el IUHIBA definió mantener la instancia de entrevistas virtuales para favorecer la accesibilidad a postulantes de todo el país. Por lo cual continuamos desarrollando las MME virtuales (MMEv) y las residencias que participaron en 2022 fueron: Cirugía General, Ginecología quirúrgica, Medicina Familiar y Tocoginecología. En estos últimos años hay experiencias de otros países publicadas acerca del diseño y puesta en marcha de MMEv para la selección de ingresantes a escuelas de Medicina, programas de residencia y post residencia. Reportan que ha sido una experiencia factible y exitosa, con el beneficio de reducción de costos (traslado, locación, etc.) y remarcan la importancia de la comunicación y preparación previa de todos los actores intervinientes (Cleland et al, 2020; Melro et al., 2022; Singh et al., 2021; Ungtrakul, 2020).

Planteo del problema y justificación de la investigación

Algunos interrogantes dentro del estudio de las MME son ¿qué miden realmente? ¿Se pueden diferenciar atributos? ¿Qué tienen en cuenta los evaluadores para las puntuaciones?. En conexión con estas preguntas, un área de interés en el análisis psicométrico de las MME hace foco en el número de factores o variables (atributos no cognitivos) que evalúa. La discusión es si se debe considerar que se evalúa un único factor o es multifactorial (Oliver et al., 2014; Sebok et al., 2014; Zaidi et al., 2016). Existe una gran variabilidad en qué atributos no cognitivos evalúa cada institución a través de las MME (Reiter & Eva, 2018). Una publicación que reporta resultados obtenidos a partir del análisis factorial (AF) sugiere que los

puntajes colapsan en un modelo de 2 factores (comunicación oral y evaluación de problemas) explicando de esta forma la varianza en su MME, aunque los dos constructos estaban altamente correlacionados en 0,87 (Oliver et al., 2014; Sebok et al., 2014; Zaidi et al., 2016). En otro estudio realizaron un análisis de tres factores, uno en cada uno de los dominios puntuados dentro de las estaciones: atributos no cognitivos, habilidades de comunicación y habilidades de pensamiento crítico (Hecker et al., 2009). Los puntajes de los atributos no cognitivos se combinaron con el promedio de calificaciones y la edad y dieron como resultado una solución de tres factores: "valores éticos y morales", "capacidad interpersonal" y "capacidad académica". Los puntajes de habilidades de comunicación cargan todos en un factor, y el pensamiento crítico carga en dos factores de pensamiento crítico separados, lo que sugiere un total de seis factores. Lemay (Lemay et al., 2007) publicó que el análisis de su MME reveló una solución de 10 factores, con cada estación formando un solo factor. Por otro lado estudios más recientes como el de la Universidad de Ciencias de Malasia, donde realizaron un análisis factorial confirmatorio obtuvieron un constructo unidimensional que les proporcionó apoyo para utilizar la puntuación media como rendimiento general (Abdul Rahim & Yusoff, 2016). Frente a las discrepancias en este área es que realizaremos un análisis factorial exploratorio (AFE) con los datos obtenidos de las MMEv realizadas en el IUHIBA en 2022, así comparar nuestros resultados con los publicados en la bibliografía internacional hasta el momento. A su vez, para contextualizar las puntuaciones y determinar qué percibían como relevante los evaluadores resulta de utilidad analizar comentarios realizados por ellos acerca de los postulantes, tal como se realizó en las MME para la selección de ingresantes a Medicina de dos universidades de Estados Unidos (Manuel et al., 2019). Este estudio analizó solo comentarios escritos y reportó que los evaluadores se centraban en siete dominios diferentes: perspectiva, presentación, cualidades, coherencia, comprensión, comunicación verbal y no verbal.

En resumen, en esta tesis me planteo obtener evidencia a nivel local (Hospital Italiano de Buenos Aires, Argentina) para estudiar el interrogante planteado en relación a si verdaderamente la evaluación a través de las MME permite diferenciar entre los distintos atributos no cognitivos preestablecidos. Por otra parte, analizar la confiabilidad, validez y aceptabilidad del método en nuestro

medio, y describir qué pensamientos y percepciones tienen los evaluadores frente al desempeño de los postulantes en su estación. Para este fin utilizaré los datos obtenidos en las MMEv realizadas en 2022.

Objetivos de la investigación

Objetivo primario

Evaluar si la estructura de los factores luego de la aplicación de las MME se corresponde con los atributos no cognitivos preestablecidos en su diseño inicial.

Objetivos secundarios

1. Explorar los pensamientos y percepciones de los evaluadores frente a las actitudes de los postulantes que modelan el proceso de selección.
2. Analizar la validez y confiabilidad de las MMEv.
3. Valorar la aceptabilidad de las MMEv de evaluadores y postulantes.

Metodología

Tipo de estudio, diseño y recolección de datos

El presente es un estudio descriptivo, analítico, transversal. Se utilizó un enfoque mixto: cuanti-cualitativo para responder a los objetivos planteados. Para el análisis cuantitativo se utilizó el registro de puntajes de las MMEv a partir de una base de datos anonimizada. Para el análisis cualitativo las estrategias para la recolección de datos fueron conversacionales (entrevistas semiestructuradas a evaluadores) y escritas (en fichas de evaluación con el fin de identificar cuáles fueron los elementos en los que se basaron para puntuar a los postulantes). Para valorar la aceptabilidad se utilizó un cuestionario online autoadministrado.

Población y ámbito del estudio

Por un lado, participaron de las MMEv postulantes que aprobaron la instancia de examen escrito y accedieron a entrevista individual en el proceso de selección de residentes en el HIBA en el año 2022, a las residencias de:

- Cirugía General (C)
- Medicina Familiar y Comunitaria (MF)
- Tocoginecología, sedes: central y San Justo, + Ginecología quirúrgica (G): son 3 residencias que forman parte del mismo servicio por lo cual se tomaron como un único grupo o especialidad.

Por otro lado, se convocó a profesionales del HIBA que participaron en el rol de evaluadores en las MMEv. Para la realización de las entrevistas el muestreo fue no probabilístico por conveniencia, es decir implica la selección de casos que representen variabilidad de diversos supuestos, y por cuotas en pos de realizar una representación de la heterogeneidad existente en los perfiles de los evaluadores, en torno a: especialidad, antigüedad y experiencia en el proceso de selección de residentes. El tamaño de la muestra (cantidad de entrevistas) se definió con un mínimo de sujetos participantes de diferente perfil debido al muestreo intencional por cuotas, y un máximo delimitado por la saturación de los contenidos investigados, por lo que se detuvo el proceso de recogida de información al constatarlo (Kornblit & Beltramino, 2004; Maxwell, 2022; Strauss & Corbin, 2002).

Desarrollo e implementación de las Múltiples Mini Entrevistas virtuales

El diseño se realizó tomando como referencia las publicaciones de Eva. et al (Eva et al., 2004, 2019) y se realizaron de manera virtual a través de la plataforma Zoom®. Cabe aclarar que el contenido de esta tesis no tuvo implicancias en el proceso de ingreso a las residencias, ya que se realizó de manera retrospectiva, una vez finalizada la selección e ingreso de residentes.

Debido al número de postulantes a entrevistar se realizaron dos circuitos consecutivos idénticos para cada especialidad médica, por decisión de la jefatura y coordinación de las residencias involucradas. Cada circuito estuvo formado por una serie de entrevistas o estaciones de formato flexible (entrevista tradicional,

escenarios hipotéticos, basados en la experiencia, interacción con otra persona, etc.) adaptadas para ser realizadas de manera virtual. Se valoraron los atributos no cognitivos determinados como relevantes para el perfil del residente (razonamiento ético, trabajo en equipo, argumentación, motivación hacia la especialidad y recepción de feedback), de acuerdo a la tabla de especificaciones (anexo I) avalada por los coordinadores de las residencias participantes. Cada circuito contó con 12 estaciones (3 con actores, similar a pacientes simulados), excepto para tocoginecología-ginecología que fueron 14 estaciones (se sumaron 2 más por contar con mayor cantidad de postulantes). Para asegurar la confidencialidad del contenido de las estaciones los participantes de la primera y segunda ronda no tuvieron contacto entre ellos. Participó un evaluador por estación que completó una ficha de evaluación (por seguridad utilizamos doble registro: en papel y online) para cada postulante con una escala de Likert para el atributo no cognitivo evaluado y el desempeño global (anexo II) y un espacio de texto libre optativo para ampliar por escrito por qué le asignó ese puntaje en el ítem de desempeño global (con el fin de contextualizar la calificación).

Para la instrucción de los evaluadores se realizó una reunión general donde se les brindó una capacitación sobre las MME y el uso del Zoom®, y 48 hs antes de las entrevistas se realizaron encuentros virtuales para entrenamiento individualizado. A su vez, actores y actrices participaron de un entrenamiento y ensayo de las estaciones de manera virtual. La organización general, entrenamientos y logística fue liderada por la investigadora principal de esta tesis junto a un equipo docente, un especialista en informática y personal administrativo del IUHIBA.

El día de las entrevistas, tanto postulantes como evaluadores accedieron desde sus domicilios a la sala virtual coordinada por el equipo docente organizador que se encontraba en oficinas del HIBA. Se utilizaron 3 cuentas de Zoom® (una por especialidad) donde el anfitrión era un docente del equipo coordinador de la residencia. En primera instancia se distribuyó a cada entrevistador a una sala privada de Zoom® (opción salas para grupos pequeños o breakout rooms), de esta manera se crearon salas nombradas como Estación 1, 2, 3, etc. Los evaluadores fueron renombrados también con el número de la estación correspondiente y

quedaban fijos en ellas. Luego se permitió el ingreso a postulantes, se les realizó un chequeo de la identidad con los datos obtenidos de la documentación brindada por el IUHIBA y fueron renombrados con un número para garantizar el anonimato durante la evaluación. Al iniciar el circuito se los asignó cada uno a una sala o estación virtual, pasados 6 minutos se enviaba un mensaje avisando que faltaba 1 min para el cambio de estación. A los 7 minutos quien tenía rol de anfitrión se encargaba de hacer la rotación de los postulantes para que pasaran a la siguiente estación. Cada vez que un nuevo postulante ingresaba a una estación el entrevistador copiaba en el chat de Zoom® las instrucciones correspondientes. Esta dinámica se repitió hasta finalizar el circuito.

El instrumento de registro (fichas de evaluación) se desarrolló en formato digital a través de un sistema basado en Google Drive, lo cual permitió obtener los rankings el mismo día de las entrevistas.

En el diseño de cada estación se definieron por consenso banderas rojas, actitudes o dichos que resultaban inaceptables para la situación planteada (ej. maltrato del actor), que implicaban como penalización que el postulante bajara un lugar en el ranking final.

Al finalizar la jornada se envió por mail un cuestionario para valorar la aceptabilidad del proceso, de carácter anónimo y voluntario a todos los participantes (postulantes y evaluadores).

Análisis

Objetivo primario: Análisis factorial exploratorio

Se exploró la validez del constructo evaluando la estructura factorial (unifactorialidad vs. multifactorialidad) de los aspectos o variables (razonamiento ético, trabajo en equipo, argumentación, motivación hacia la especialidad y recepción de feedback) en un análisis factorial exploratorio (AFE). Se tuvo en cuenta en primera medida el Criterio de Kaiser programado por defecto en los programas de computación como el SPSS y el STATA. En este sentido se consideraron solamente los factores que en el primer análisis (antes de las rotaciones) tengan un eigenvalue (o valor propio) mayor de 1 (Kaiser, 1970). Por otro lado, se tuvo en

cuenta el criterio de Carmines que plantea que si la varianza explicada por el primer factor es de 40% o más, la estructura es unifactorial (Carmines et al., 1979). También consideramos el criterio de Lord que plantea que el cociente entre la diferencia del primer eigenvalue y el segundo eigenvalue, y la diferencia entre el segundo y tercer eigenvalue debe ser importante, con un valor mayor a 3 (Hattie, 1985; León, 2006). Y por último se evaluó el criterio de Gorsuch, que consiste en tomar en cuenta sólo los factores que tengan un eigenvalue mayor de 1,41 (Gorsuch, 1997).

También se realizó una validación de criterio externo o concurrente entre el factor “Valoración” (razonamiento ético, trabajo en equipo, argumentación, motivación hacia la especialidad y recepción de feedback) de la herramienta MME y la variable “global” de la misma. Nuestra hipótesis consistió en pretender una correlación (Pearson) cercana a 0,7/0,8 (y como mínimo 0,4) entre el puntaje del factor compuesto por las siguientes variables: razonamiento ético, trabajo en equipo, argumentación, motivación hacia la especialidad y recepción de feedback; y la variable “global”, por entender que se trata de instrumentos que miden constructos similares, como señala la bibliografía (Streinmer, 1993). Se consideró para determinar dicho factor, guardar previamente como variable las “puntuaciones factoriales” del mismo (método Regresión, SPSS), para luego realizar correlación Pearson entre esta última variable (puntuaciones factoriales del factor) y la variable “global”.

Cálculo del tamaño muestral: Para realizar el AFE y la evaluación de la consistencia interna fue necesario contar con un mínimo de cinco observaciones por variable, si el n fuera mayor de 100, o diez observaciones si el n fuera menor de 100, según lo exige la literatura especializada (Streiner, 1994), y nunca menos de 50 como valor absoluto (Hair et al., 2018).

Objetivos secundarios:

- **Análisis de contenido:** El trabajo de campo se llevó a cabo en septiembre de 2022, inmediatamente posterior a la ocurrencia de las MMEv. Los datos se obtuvieron de dos fuentes: *comentarios escritos* por evaluadores en relación a cuáles fueron los elementos en los que se basaron para definir los puntajes y

entrevistas a evaluadores. Se utilizó la técnica de entrevistas semiestructuradas, utilizando una guía de preguntas ad-hoc (anexo III), cuya flexibilidad ayuda a abordar mejor el fenómeno. Las mismas fueron individuales y duraron entre 30-45 minutos. El material fue audio-grabado y luego se realizó la desgrabación textual a través de un servicio de transcripción que garantizó que todos los datos identificatorios fueran anonimizados previo a la etapa de análisis.

Se utilizó como marco teórico el análisis temático (Braun & Clarke, 2006). El análisis de la información se efectuó por dos investigadoras (la autora y la co-directora de esta tesis), para lo cual se utilizó el software Atlas.ti (versión 22). Se organizaron los documentos con los comentarios escritos y los registros de entrevistas transcritas en archivos Word.

Se siguieron los pasos descritos por Braun y Clarke para el análisis temático (Braun & Clarke, 2006): familiarización con los datos, generar códigos iniciales, generar temas a partir de los códigos, revisión de los temas, definir y nombrar temas (mapa temático o sistema de códigos final) y producir el reporte final. Más allá de la distinción en pasos, el análisis empezó ya desde el inicio (incluso durante la recolección de datos) e implicó un movimiento constante hacia adelante y atrás entre todo el conjunto de datos y los extractos codificados. La escritura también se inició en los primeros pasos al anotar ideas y posibles esquemas de codificación que se continuaron durante todo el proceso.

En primer lugar, ambas investigadoras leyeron de manera independiente los documentos repetidamente para obtener una perspectiva global y familiarizarse con los datos. Luego cada investigadora comenzó la codificación de manera separada, resaltando palabras, oraciones y frases para organizarlas en grupos significativos, mientras tomaban notas aparte haciendo comentarios sobre posibles temas emergentes. Luego cada una revisó sus propios códigos para identificar patrones y llegar a un listado inicial de temas. Se intercambiaron este listado inicial para una revisión previa a la primera reunión, en la cual de manera conjunta revisaron el listado de temas y códigos conciliando diferencias, marcando superposiciones y armando un primer mapa temático (similar al sistema de códigos de la teoría fundamentada). Se continuó el análisis conjunto definiendo específicamente cada tema, de esta manera se llegó a la confección del mapa temático (Braun & Clarke,

2006). Se construyó teniendo en cuenta tanto las preguntas centrales de la investigación como los hallazgos emergentes desde el contenido.

Finalmente, se elaboró un informe narrativo con los resultados encontrados que comunica la información sobre cada tema, presenta el sistema de códigos utilizados y es ilustrado con una selección de citas textuales.

- **Confiabilidad**

Se evaluó la consistencia interna utilizando el análisis del coeficiente alfa de Cronbach. Este coeficiente mide la consistencia interna de los ítems de un instrumento (grado en que co-varían), en este caso cada estación representa un ítem de análisis y el instrumento de evaluación las MME. Este método de correlación entre ítems puede ser utilizado ya que se utiliza una escala de Likert, en especial con escalas como esta de 10 ítems que pueden considerarse continuas. Este coeficiente varía entre 0 y 1, siendo 1 el valor asignado a la máxima confiabilidad del instrumento y deriva de la Teoría Clásica de los Tests y permite estimar el error global de medición (Zúñiga-Brenes & Rojas, 2011).

Se utilizó la teoría de la generalizabilidad (teoría G) para determinar la varianza de los postulantes y la capacidad de generalización de los resultados. Esta teoría utiliza el análisis de la varianza (ANOVA) que permite medir la confiabilidad de un instrumento de evaluación al identificar y cuantificar la importancia de cada una de sus fuentes de variación. El error aleatorio para la Teoría Clásica de los Tests en el contexto de la teoría G se refiere a diversas fuentes no explicadas de variación y se denominan "facetas" (Zúñiga Brenes & Rojas, 2011). Se calculó el coeficiente G y el porcentaje que aporta la faceta de diferenciación (objeto de medida: postulantes) a la varianza total a través del software EduG 6.1e (Generalizability Study. Working Group - Edumetrics - Quality of measurement in education of the Swiss Society for Research in Education). Para todos los análisis se considerará un coeficiente satisfactorio igual o mayor a 0.70 (Pau et al., 2013).

- **Validez**

La definición tradicional afirma que un instrumento es válido si mide lo que se pretende medir (Nunnally, 1995). Sin embargo a partir de las publicaciones de

Messick (Messick, 1989, 1995) se adopta una concepción más moderna, así al validez se refiere a la evidencia presentada para apoyar o refutar la interpretación asignada a los resultados de la evaluación, de esta manera siempre es elaborada como una hipótesis y luego se recopilan los datos para apoyarla (Downing, 2003). Desde esta perspectiva ya no se hace referencia a distintos tipos de validez sino a la validez como un concepto unitario y se habla de un proceso de recolección de diferentes tipos de evidencia como contribuyentes a la validez de constructo (Messick, 1989). Este marco de referencia se incorpora en 1999 en los "Standards of Educational and Psychological Measurement" (AERA (American Educational Research Association), APA (American Psychological Association) & NCME (National Council on Measurement in Education), 1999) proponiendo 5 tipos de evidencia de validez, basadas en: el contenido, estructura interna, proceso de respuesta, relaciones con otras variables externas y las consecuencias de la aplicación del test (Downing, 2003). De esta manera para verificar la validez de las inferencias realizadas a partir de las puntuaciones de un test se recolecta evidencia que confirme o refute esas inferencias (Eignor, 2013).

- Aceptabilidad

La aceptabilidad de un instrumento habla acerca de cuán adecuados los evaluados y los evaluadores, consideran los métodos de evaluación propuestos. En este caso, para valorar la aceptabilidad se envió un cuestionario online, inmediatamente al finalizar las MME, a postulantes y evaluadores. Para minimizar el riesgo de sesgo de cortesía, mantener la privacidad y la confidencialidad en las respuestas, el cuestionario fue autoadministrado y anónimo, sin presentar preguntas que permitan la identificación del sujeto. Para responder a cada enunciado se utilizó una escala tipo Likert con 5 puntos (1: totalmente en desacuerdo / 3: ni de acuerdo ni en desacuerdo / 5: totalmente de acuerdo). El cuestionario fue diseñado tomando como base publicaciones existentes sobre el tema y para indagar áreas de interés para la institución (Hofmeister, Lockyer, & Crutcher, 2008) (anexo IV).

Consideraciones éticas

Todos los participantes convocados para responder cuestionarios y entrevistas dieron su consentimiento informado oral y su decisión voluntaria, para garantizar la confidencialidad de los informantes se utilizaron etiquetas durante el almacenamiento, y el plan de trabajo cuenta con aprobación por el comité de ética institucional (CEPIU #0016-22) (anexo V).

Resultados

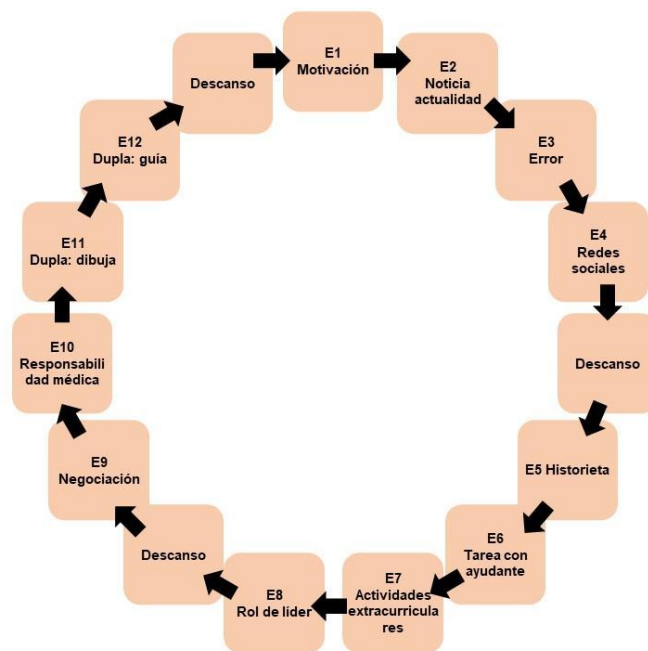
En 2022 se realizó el proceso de selección de residentes a través de la metodología de MMEv para las especialidades de C, MF y G. Participaron en total 76 postulantes, el 18,4% registraba género masculino y el 84,2% femenino en su documento de identidad, la media de edad fue de 27 años (Tabla 1). Como se menciona en el apartado de desarrollo de las MMEv, para cada especialidad se realizaron circuitos separados y en paralelo con las demás residencias. Cada circuito contó con las mismas 12 estaciones (2 estaciones extras para G) + estaciones de descanso, de 7 minutos de duración cada una (ver figura 1). Participaron 38 evaluadores, casi en su totalidad fueron médicos docentes de la especialidad involucrada. En MF participó una docente de otra disciplina (psicología) y en G una magíster en educación de otra especialidad médica. Se utilizó el software de videollamada Zoom® con 3 cuentas propiedad del IUHIBA.

Tabla 1. Datos demográficos de postulantes

	Total	Cirugía	Medicina Familiar	Tocoginecología		Ginecología
				Sede Central	Sede San Justo	
Vacantes disponibles	16	5	5	4	1	1
Postulantes que asistieron/citados n/n total (%)	76/88 (86,4)	23/25 (92)	17/25 (68)	20/20 (100)	10/10 (100)	6/8 (75)
Edad, en años media (DS)	27,1 (5,3)	25,2 (0,7)	30,8 (9,7)	25,6 (0,9)	28,4 (1,9)	26 (1,7)
Género según documento n (%)						
Masculino	14 (18,4)	6 (26,1)	4 (23,5)	3 (15)	1 (10)	0 (0)

Femenino	64 (84,2)	17 (73,9)	13 (76,5)	17 (85)	9 (90)	8/8 (100)
----------	-----------	-----------	-----------	---------	--------	-----------

Figura 1. Modelo del esquema de la rotación a través del circuito



E: estación

El análisis descriptivo de las MME se presenta en la tabla 2. Las puntuaciones medias de las estaciones oscilaron entre 5,37 y 7,87, siendo 10 el máximo puntaje posible por estación y la media del total fue de 80,28 con un desvío estándar de 16,40.

Tabla 2. Media y desvío estándar de las puntuaciones por estación

Estación	Media (DS)	Mínimo	Máximo
E1 - Motivación	7,57 (1,95)	1,40	10
E2 - Noticia de actualidad	5,63 (3,31)	1	10

E3 - Error	7,09 (2,32)	1,20	10
E4 - Redes sociales	7,35 (2,44)	2	10
E5 - Historieta	6,75 (3,01)	1	10
E6 - Tarea con ayudante	7,29 (2,48)	1	10
E7 - Act. extracurriculares	5,37 (2,95)	1	10
E8 - Rol de líder	7,87 (3,17)	1	10
E9 - Negociación	6,32 (2,74)	1,80	10
E10 - Responsabilidad médica	7,04 (2,40)	2	10
E11 - Dupla: dibuja	6,21 (2,65)	1	10
E12 - Dupla: guía	5,68 (3,16)	1	10
Total	80,28 (16,40)		

Objetivo primario: Análisis Factorial Exploratorio

Para responder al objetivo primario se realizó el AFE, utilizando el programa IBM SPSS Statistics v21. Este método sirve para agrupar variables o factores que correlacionan fuertemente entre sí y verificar si es posible describir un constructo uni o multifactorial medido en las MME.

La medida de adecuación muestral de Kaiser-Meyer-Olkin tuvo un resultado óptimo de 0,735, y la prueba de esfericidad de Bartlett fue claramente significativa (0,0001).

Para determinar la validez del constructo, y certificar la unidimensionalidad (unifactorialidad) de la herramienta, se realizó un AFE (Ver Tablas 3 y 4).

Tabla 3. Varianza total explicada

Componentes	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	2,494	49,875	49,875	2,494	49,875	49,875
2	0,902	18,046	67,920			

3	0,706	14,116	82,036			
4	0,464	9,272	91,308			
5	0,435	8,692	100,000			

Tabla 4. Matriz de los componentes

Variables	1 componente extraído
Motivación	0,772
Argumentación	0,708
Trabajo en equipo	0,752
Razonamiento ético	0,611
Recepción de feedback	0,677

Se cumplió el criterio de Kaiser, que exige que se consideren sólo los factores que presentan un eigenvalue o autovalor mayor de 1, ya que el único autovalor mayor de 1 fue el primero (con un valor de 2,494), alcanzando el segundo 0,902.

Se verificó también el criterio de Carmines, que establece una estructura unifactorial ante una varianza explicada por el primer factor superior a 40%: en nuestro caso fue 49,875%.

Se evaluó también el criterio de Lord, que estudia el cociente entre la diferencia del primer autovalor y el segundo ($2,494 - 0,902 = 1,592$), y la diferencia del segundo y tercer autovalor ($0,902 - 0,706 = 0,196$). Este cociente ($1,592/0,196 = 8,122$) arrojó resultados altamente satisfactorios, ya que debía ser mayor de 3.

Por último se tuvo en cuenta el criterio de Gorsuch, comprobándose que un único factor presentaba un eigenvalue mayor de 1,41, concretamente 2,494, y el segundo apenas 0,902, reforzando nuevamente la idea de unifactorialidad.

Con respecto de la validez concurrente o de criterio externo, se obtuvo un resultado óptimo de coeficiente Pearson de 0,960.

La confiabilidad fue evaluada a través del coeficiente Alpha de Cronbach, que arrojó un valor fuerte de 0,743.

Objetivos secundarios

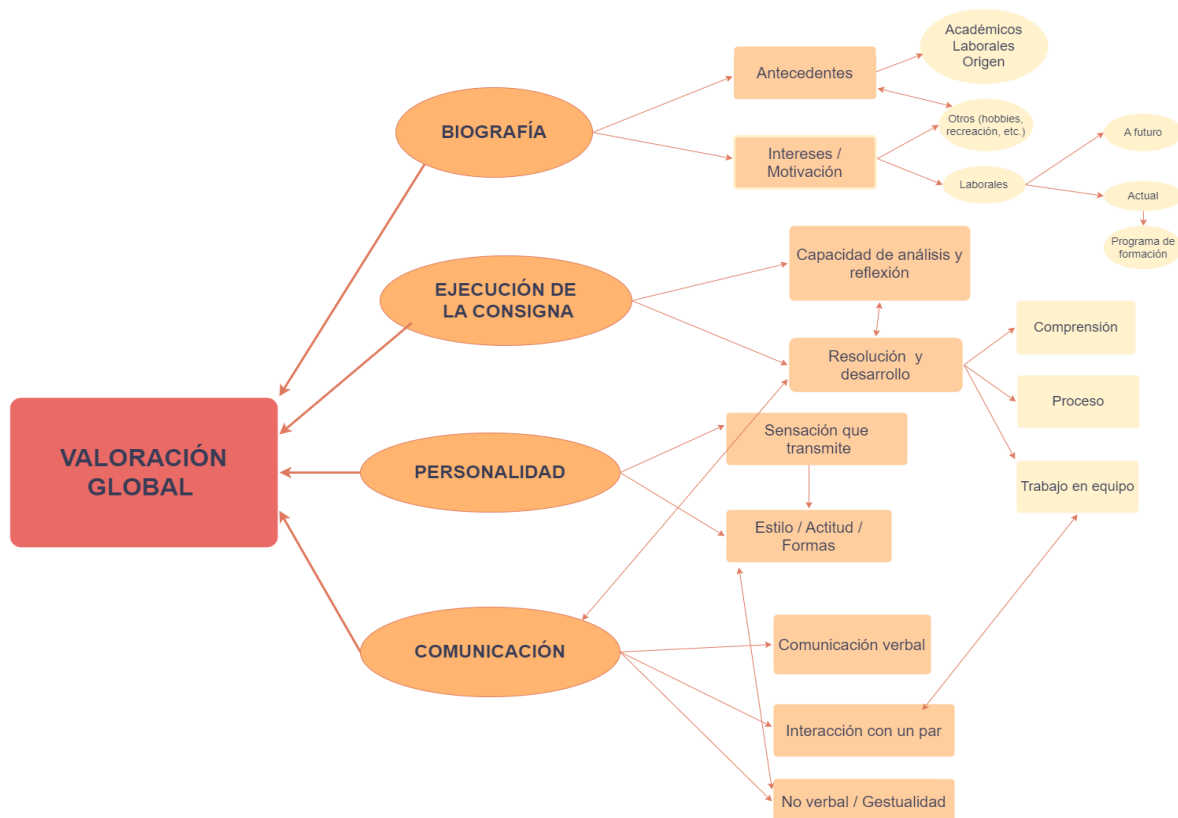
Análisis del contenido

Se utilizó el enfoque cualitativo a través del análisis de contenido temático para responder al primer objetivo secundario planteado. Para ello el corpus se obtuvo a partir de entrevistas y comentarios escritos por los evaluadores, con el fin de que describan e identifiquen cómo construyeron la valoración del desempeño (ej. en qué elementos, actitudes se basaron para puntuar).

Se entrevistaron 8 evaluadores, el 50% se identificó con el género masculino y 50% con el femenino, la media de edad en años fue de 46,2 (DS 9,41), el 37,5% participaba por primera vez en las MME. Por otro lado, se recolectaron 592 comentarios, lo que representa que un 60,2% de las fichas de evaluación tenían algún comentario escrito.

A continuación, se presenta el mapa temático (figura 2), con sus temas, subtemas y relaciones, definidos para el análisis de las entrevistas y comentarios escritos, luego se definen y ejemplifican mediante citas.

Figura 2. Mapa temático



En función a los objetivos del estudio se identificó que los códigos podrían agruparse en cuatro temas principales y que todos ellos confluyen en una dimensión que es la *Valoración Global* del postulante:

1. Biografía
2. Ejecución de la consigna
3. Personalidad
4. Comunicación

En el siguiente apartado se da una descripción abreviada de qué representaba para las investigadoras cada uno de los temas y subtemas.

La valoración global como dimensión que abarca a los demás temas la relacionamos con la apreciación general que hacen los evaluadores de los postulantes y su desempeño en la estación. Surge de comentarios estrechamente relacionados con los cuatro temas principales y dependiente de qué "mira" o que le resulta relevante al evaluador por fuera de los objetivos de la estación que le fue

asignada. El primer tema hace referencia directamente a datos biográficos y del currículum del postulante, solo evaluados en las dos estaciones con formato de entrevistas tradicionales. Mientras que los siguientes temas nos permiten analizar comentarios acerca del desempeño en el resto de las estaciones enfocando en diferentes aspectos como la ejecución, comunicación e intercambio y estilos.

Definición de códigos

1. Biografía: comentarios relacionados a aspectos de la historia personal y experiencia de los postulantes.

a. Antecedentes: recorrido académico y laboral, orígenes, donde vive/vivió, hobbies, etc.

b. Intereses: en lo que respecta a la residencia de la especialidad elegida, a la institución, a su futuro laboral y extra laboral, recreación.

2. Ejecución de la consigna: comentarios con respecto a cómo es el desarrollo o cumplimiento de la estación en relación a la situación planteada; ya sea la descripción literal de lo que hace o relacionado al nivel de comprensión, análisis y reflexión que demuestra el postulante.

a. Resolución: referente a pautas del desarrollo de la estación en relación a la consigna.

i. Comprensión: entendimiento de lo que se le pedía que hiciera en esa estación.

ii. Proceso: forma en la que resuelve la consigna, descripción de lo que hace, adjetivación de la manera en la que lo hace.

iii. Trabajo en equipo: manera en la que resuelve la consigna junto a otro en vivo o en un supuesto.

b. Capacidad de análisis y reflexión: comentarios que demuestran un proceso más profundo al resolver la consigna (no solo si hace o no hace, comprende

o no comprende). Por ej. identificación de problemas, priorización, propuesta de alternativas, razonamiento ético, creatividad, etc.

3. Personalidad: comentarios que describen aspectos relacionados con la forma de presentarse, mostrarse, desenvolverse, dirigirse y a la vez lo que esto produce en el otro.

- a. Sensación que transmite: impresión general o química que genera en el evaluador.
- b. Actitud: relacionado con un estilo o postura de conducirse durante la entrevista, tono y modo de transmitir.

4. Comunicación: comentarios relacionados a la transmisión de información en el intercambio. En las MME dependiendo de la estación puede ser directo con el evaluador o con un tercero (actor, colega) y en este caso el evaluador actúa solo como observador.

- a. Comunicación verbal: en referencia a lo que decían a través de las palabras.
- b. Interacción con un par: se le asigna este código a aquellos comentarios que describen la manera en la que interacciona en vivo con otro postulante o actor en rol de par. Diferenciándose de la interacción con el evaluador que está en otro nivel de poder.
- c. No verbal: relacionada con lo expresado a través de gestos, posturas, imagen.

En la tabla 5 se exponen algunas citas textuales de los evaluadores, tanto de entrevistas como de comentarios escritos, a modo de ejemplificar la definición de códigos y temas.

Tabla 5. Citas de evaluadores organizado por temas.

Temas principales	Subtemas		Comentarios ^a
BIOGRAFÍA	Antecedentes		<i>Llama la atención que estudió comunicación social, luego Medicina, hizo algunos meses de residencia de gineco, luego se pasó a residencia de clínica, a los 2 años dejó (...) CE19</i>

	Intereses	<p><i>Mucha actividad laboral, conoce el HIBA. Vino a visitar el Servicio. Le gusta la docencia. CE35</i></p> <p><i>Tiene proyecto de hacer posbásica al finalizar la resi. CE23</i></p>	
EJECUCIÓN DE LA CONSIGNA	Resolución	Comprensión	<p><i>Al que no puede contestar la consigna también le bajaba puntos, me parece importante que pueda leer el texto y entenderlo. E1</i></p> <p><i>No entendió la consigna, no supo conducir al compañero, no mostró personalidad, mucho silencio. CE33</i></p>
		Proceso	<p><i>Escucha a su amigo para ver en qué puede ayudarlo. Cree que lo más importante es que vea a su médico de cabecera. Lo orientaría, sin dar diagnósticos. (...) Resalta como fundamental el aval del padre para que su hijo le muestre los estudios. Trata de que esté presente el padre en la devolución de los estudios. CE27</i></p> <p><i>Negocia pero no resuelve del todo, queda en buscar algo (...) falta de iniciativa y resolución. CE36</i></p>
		Trabajo en equipo	<p><i>La otra postulante le dijo ¿te puedo preguntar algo? Y le respondió No, yo te voy a decir lo que tenes que hacer. (...) la consigna dice que justamente te pueden preguntar y vos contestar (...). E6</i></p> <p><i>Mostró aptitudes para trabajar en grupo, buena recepción a críticas. Además por cómo guió parece ser una persona práctica y eso es una buena característica para un residente. CE33</i></p>
	Capacidad de análisis y reflexión	<p><i>(...) a mi la inteligencia me sorprende más allá que no estoy evaluando inteligencia, cuando te das cuenta que te estoy llevando a un lugar y el tipo sale y siempre me deja en offside. E1</i></p> <p><i>Las responsabilidades del residente al principio son muy limitadas, y lo que vemos como fundamental es que tengan noción de la limitación que tienen (...) Por eso me pareció muy valioso, los que realmente se dieron cuenta que tenían que contactar a un superior. E7</i></p> <p><i>Identifica el problema de privacidad/intimidad. Plantea una solución reflexiva. Empática. Busca abordar la dificultad que existe entre ella y el implicado. Detecta los posibles perjuicios del implicado. Define en forma adecuada el uso responsable de las redes sociales. CE26</i></p>	
PERSONALIDAD	Sensación que transmite	<p><i>En la mitad de la estación, ya te das cuenta como lo quieres puntuar (...) los 7 minutos alcanzaban para hacer todo lo más bien. E6</i></p> <p><i>Que te genere algo, y efectivamente una chica (...) a pesar de haber dado un excelente examen (de conocimiento), fanática de la cirugía y haber tenido todo para entrar, ha habido algo en sus entrevistas que nos generó a todos los mismos. (...) es mucho la presencia y lo que uno impone también. E2</i></p>	
	Actitud	<p><i>Lo primero que hizo fue decir que ella "no era buena para este tipo de actividades" (...) siempre poniéndose en desventaja con respecto a su oponente. CE32</i></p> <p><i>Espontánea. Genuina. A pesar de la propuesta abrumadora supo negociar con calidez y creatividad. Excelente postulante!!!! CE24</i></p> <p><i>El que demuestra ser un canchero, no cae muy bien tampoco. Yo creo que la timidez juega en contra, pero también ser demasiado extrovertido, en el sentido de ser un confanzudo. E2</i></p>	

COMUNICACIÓN	Verbal	<p><i>El candidato que arranca y te dice todo de una sin ningún tipo de freno, sin ninguna traba, con un buen vocabulario te compró y punto, ya no hay mucho más. E3</i></p> <p><i>Lo segundo que valoro es el orden para comunicar lo que querían comunicar, muchos empiezan a hablar y pierden el hilo de lo que estaban diciendo. E7</i></p> <p><i>Muchas palabras sin contenido real. CE22</i></p>
	Interacción con un par	<p><i>Parece poco receptiva a la crítica y poco abierta a aceptar su rol en un resultado adverso. Por momentos es autoritaria en su forma de responder, interrumpe mucho y no escucha demasiado. Siempre marca el error ajeno antes de reflexionar sobre su accionar. CE32</i></p> <p><i>Era re importante que se saludaran, el tono que voz que usaban, un tono amable agradable o si era rígido y autoritario, eso también vale. (...) en un ambiente más de compañerismo digamos o de sos mi contrincante, se notaba mucho en algunas parejas. E6</i></p>
	No verbal	<p><i>No podía ser actuado, la forma de sentarse, la forma de decir las cosas (...) Entonces estas cosas de poder ser genuinos, espontáneos, sinceros, no enojados, la forma de hablar, todo eso lo tuve en cuenta a la hora de puntuar. E8</i></p> <p><i>Si vos estás viendo que todos los chicos se presentan bien vestidos, bien peinados, y viene uno que está con el pelo (hace gesto de alborotado), (...) yo ya le pierdo un poco la objetividad a esa evaluación. E7</i></p>

^a E: entrevistado, CE: comentario escrito, cada evaluador se identificó con un número asignado al azar.

Dimensiones analíticas

De los relatos y comentarios surgen dos dimensiones analíticas centrales en las cuales los evaluadores se enfocan y que contribuyen de manera similar al momento de definir la puntuación: la primera impresión y la manera en la que resuelven la tarea. Además de estas surgen en segundo plano otras dimensiones, no por eso menos importantes, una en referencia a conocer los antecedentes académicos y otra a la dificultad de evaluar diferenciadamente los puntajes definidos para la estación.

La **primera impresión**, esa valoración emitida rápidamente sobre otros, la cual se basa en varios aspectos de la personalidad (actitud, imagen, postura) y comunicación (verbal y no verbal: saludo, presentación, tono).

La simpatía, como te mira, como te saluda, la locuacidad, la conexión, viste vos conectas con una persona más rápido que con otra y eso es inconsciente, yo intentaba salir de esa historia, pero te pasa en todas las entrevistas. E1

Ya desde el saludo para mí marcaban diferencia, el que saludaba, el que hablaba bien, el que tenía buena presentación. Creo que no es algo a lo que apunta la entrevista, pero somos humanos. (...) Creo que es lo que a uno le transmite también, que ya lo detectas. E2

Al inicio ya poniendo buena predisposición (...) ya le ponían esa actitud positiva, yo la sentía, era espontánea, genuina (...). E8

Como refieren E1, E2 y E8 ya al inicio de la entrevista se genera una percepción que les resulta difícil de esquivar, lo cual no se relacionan con la consigna de la estación sino que surge espontáneamente por el hecho de interactuar con otra persona. E1 y E8 lo argumentan como algo que no depende de su voluntad (inconsciente, conexión, sentimiento).

En línea con esta dimensión, algunos evaluadores mencionaron que el origen, la diferencia generacional y las distintas formas de pensar o expresarse podrían sesgarlos pero lo tenían presente para evitarlo deliberadamente. Otros refirieron darle importancia a la imagen, pero que en la virtualidad resulta más dificultoso apreciarlo. De esta manera surgió que la mayoría intentaba atenerse a las instrucciones de la estación para evitar posibles sesgos intrínsecos.

Yo me propuse antes de empezar que voy a escuchar gente joven de otra generación, con modelos de pensamiento mucho más parecidos a los de mis hijos, por ejemplo, y que yo se que pensamos distinto en un montón de cosas, y dije voy a tratar de cuando escuche una cosa que me resulte distinto a mi generación, intentar no evaluarla. E1

Creo que es inevitable pensar que hay algún sesgo. Por ejemplo, sé que me juega más en la presencial, que en la virtual. A ver, no por prejuizar, pero el estilo de como venía vestido. No el vestido, si me gustaba o no, pero alguien que venía más suelta, menos acartonada, menos estricta en sus cosas me suele gustar más. Lo veo como más plástico en el buen sentido, como persona, ¿no? Que alguien que lo veo con una actitud más rígida, más formal, ese ya por ahí me simpatiza quizás un poquito menos, pero bueno trato de no juzgarlo por eso. E5

Si noto (por la tonada) que viene alguien del interior, creo que por ser del interior se rompen el alma por estar acá, compiten, les va bien. (...) Es un sesgo, aunque no siempre es un reflejo fiel de la realidad. E7

Algo que a podría ser un sesgo para mí, es si mira a los ojos, la seguridad, la tranquilidad de poder mirar a los ojos, y no como escondiéndose (...) pero en la virtualidad, es difícil porque tienen que mirar la cámara. E8

E1 hace referencia a la diferencia generacional y cómo eso puede influenciar en la evaluación del desempeño si no se tiene en cuenta de manera explícita. E7 fue el único que describió como sesgo propio que el origen puede influenciarlo de manera positiva (solo al notarlo en la tonada), haciéndolo pensar en la historia del postulante que por ser del interior del país se esforzó más para tener esta oportunidad. Mientras que E5 y E8 enfocan sus posibles sesgos en la imagen y actitud, para E5 la formalidad es un aspecto que no valora positivamente.

La segunda dimensión se relaciona con la **resolución** de la situación problemática planteada en cada estación en particular, como la capacidad de identificar y tener en cuenta múltiples niveles y perspectivas de las diferentes partes interesadas en la toma de decisiones y posición. Poder equilibrar correctamente los pros y contras de la situación.

Ósea ves esas cosas también, vos decís bueno no es solamente como da las indicaciones, como jerarquiza, como se organiza, como ordena, como expresa, qué se yo, y el otro cómo da feedback, sino que además es que no entendieron la consigna, no leyeron bien, no la repasaron. E4

Los que intentaban encontrar la vuelta ellos mismos mientras iban hablando, de encontrar como distintas opciones, sino va esto, haría esto, sino lo otro. Me gustaban más que los otros (...) que eran más correctos en sus respuestas (...) y menos de repensarse en una situación parecida. (...) que se les ocurran alternativas que realmente demuestren que en esa situación tendrían herramientas para pensar. E5

E4 entiende que además de la comprensión de la consigna se puede distinguir en cómo un postulante prioriza y se comunica con el otro. En la resolución, E5 valora la autorreflexión y que al brindar varias alternativas de accionar se reflejen herramientas de pensamiento de esa persona.

A diferencia de lo que observamos en años anteriores de MME, en ningún momento surgió que los evaluadores hayan indagado información del currículum de los postulantes si no fue específicamente indicado en su entrenamiento y tampoco lo mencionaron como desventaja o queja hacia la modalidad. Esto resulta relevante ya que lo habitual en las entrevistas panel es que cuenten con el curriculum, promedio de la carrera y nota del examen de conocimiento y es una "costumbre" que pueden acarrear muchos docentes y resultar conflictiva a la hora de ocupar el rol de

evaluador en las MME. De esta manera demostraron respetar las instrucciones brindadas en el entrenamiento y atenerse a las indicaciones de sus estaciones.

Otro punto relevante en el que acordaron los evaluadores fue la dificultad al momento de puntuar por separado el desempeño global del atributo no cognitivo específico de su estación (ej. argumentación y desempeño global).

He puesto pocas diferencias entre uno y otro puntaje, la mayoría fue igual, porque creo que estaba muy imbricado (...). E5

Mirá, en el chico que me mintió creo que diferencie desempeño global de argumentación, porque los argumentos del pibe eran buenos pero capaz que el desempeño global se lo bajé (...) después no me daba cuenta, era muy sutil, para mí eran lo mismo. E1

Surge del relato que los evaluadores consideran que ambos puntajes se encuentran íntimamente relacionados, E5 refirió que generalmente ponía la misma nota o con muy poca diferencia numérica. Como menciona E1, excepcionalmente en un único postulante pudo diferenciar el puntaje del atributo y del desempeño global.

Confiabilidad

En base a los datos obtenidos de los 76 postulantes utilizando el programa IBM SPSS Statistics v21 se evaluó la consistencia interna del instrumento con el coeficiente alfa de Cronbach, cuyo resultado fue de 0,724.

Se midió la correlación ítem/ítem total con el fin de verificar el poder discriminante de cada ítem (estación), cuando esta correlación es igual o superior a 0,2 se considera que discrimina bien y se está midiendo el mismo constructo. Al calcularla se observa que ninguna de las estaciones tiene correlación negativa (lo que afectaría significativamente la confiabilidad), y aunque las estaciones 4 y 5 tienen una correlación menor a 0,2 si las elimináramos el coeficiente Alpha de Cronbach no sufriría un cambio significativo (tabla 6).

Tabla 6. Correlación Ítem/ ítem total

Ítem (estación)	Correlación Ítem/ ítem total	Coefficiente Alpha de Cronbach con ítem (estación) eliminado
E1 - Motivación	0,516	0,694
E2 - Noticia de actualidad	0,312	0,715
E3 - Error	0,217	0,723
E4 - Redes sociales	0,185	0,727
E5 - Historieta	0,121	0,740
E6 - Tarea con ayudante	0,402	0,701
E7 - Act. extracurriculares	0,544	0,679
E8 - Rol de líder	0,352	0,708
E9 - Negociación	0,442	0,695
E10 - Responsabilidad médica	0,305	0,713
E11 - Dupla: dibuja	0,544	0,688
E12 - Dupla: guía	0,478	0,688

La tabla 7 muestra el modelo de dos facetas que se utilizó para el análisis de generalizabilidad, postulantes (P), estaciones (E) y su interacción (PxE). Se puede observar que el porcentaje aportado por la faceta de diferenciación (P) al total de la varianza fue de 16,7%. La faceta E explica el 7,2% de la varianza. En este modelo la faceta PxE se combina con el error aleatorio y juntos explican el 76,1% de la varianza del puntaje de las MME. El Coeficiente G relativo fue 0,72.

Tabla 7. Análisis de la varianza para el modelo PxE

Facetas	Grados de libertad	Media cuadrática	Componentes de varianza aleatorios	%	Error estándar
Postulantes (P)	75	22,406	1,353	16,7	0.302
Estaciones (E)	11	50,541	0,584	7,2	0.261
P x E	825	6,176	6,176	76,1	0.304

Nota. Elaborada en base a (St-Onge et al., 2014)

Validez

Como fue mencionado en la sección de metodología, al considerar la validez como un concepto unitario, donde todo es validez de constructo, se requieren diferentes fuentes de evidencia empírica para respaldarla. A continuación se detalla el proceso para la validación del instrumento que fundamentalmente implica consulta a jueces expertos, aplicación de un piloto y análisis psicométrico básico (Rojas, 2013).

En relación al contenido, se realizó el diseño de la tabla de especificaciones, que aseguró la evaluación de una muestra representativa de cada atributo no cognitivo predefinido para el perfil de residente para luego escribir las diferentes estaciones de las MME. Tanto la tabla de especificaciones como las estaciones en su totalidad fueron revisadas, comentadas y aprobadas en su versión final por expertos (en este caso en la docencia de posgrado y selección de residentes). En el caso de las estaciones donde participaron actores se redactó el libreto, se realizaron sesiones de entrenamiento y prueba equiparable al uso de pacientes simulados en un ECOE.

Con respecto al proceso de respuesta se intenta controlar o eliminar posibles fuentes de error asociadas a la administración del instrumento (Downing, 2003). Para lo cual se llevó adelante la prueba de cada estación a través de Zoom® con la finalidad de obtener la redacción definitiva (ej. claridad de las instrucciones editando palabras de difícil comprensión o ambiguas), controlar que la duración fuera acorde a la tarea solicitada, ensayar con actores, probar el desarrollo en la plataforma, etc. También se realizó una prueba piloto de la carga de datos en las fichas de evaluación online y obtención de puntuación. Con el objetivo de que los postulantes se familiarizaran con el formato de las MME virtuales, la semana previa realizamos un encuentro virtual para presentarles el método y la dinámica, dar especificaciones técnicas, realizar una prueba del uso de Zoom® y salas para grupos pequeños, responder dudas y brindar indicaciones para el día de las entrevistas. Con los evaluadores tuvimos encuentros virtuales por estación para entrenamiento, lo que implicó la lectura detallada de la estación asignada, explicación de la dinámica, revisión de aspectos técnicos del uso de Zoom® y de la carga de los puntajes. Ya que el puntaje total resulta de la suma de los puntajes de todas las estaciones es

necesario evitar la pérdida de datos que afectaría negativamente la puntuación de un postulante, para lo cual se contó con doble registro (online y papel) e inmediatamente finalizadas las MME se contactó a los evaluadores de los cuales se registró algún dato faltante.

La estructura interna como fuente de evidencia se relaciona con las características psicométricas de las MME. Previamente ya fue reportado el resultado del análisis factorial exploratorio, la consistencia interna con el coeficiente alfa de Cronbach, la confiabilidad a través de la teoría G para evaluaciones de desempeño, la generalizabilidad y estimación de diferentes fuentes que contribuyen al error en este tipo de evaluaciones.

Al contar con experiencia en la realización de MME desde 2016 y en el formato virtual por Zoom® en 2020 y 2021 muchos aspectos relacionados con el diseño, la organización e implementación, fueron siendo mejorados año tras año.

Aceptabilidad

Para valorar la aceptabilidad del método se realizó un cuestionario online. Con el fin de minimizar el riesgo de sesgo de cortesía, mantener la privacidad y confidencialidad en las respuestas, el cuestionario fue autoadministrado y anónimo. El análisis del mismo fue realizado luego de la emisión del ranking final. La tasa de respuesta fue del 84,2% para postulantes y 60,5% para evaluadores (Tabla 8).

Tabla 8. Tasa de respuesta por especialidad al cuestionario de aceptabilidad

Especialidad	Postulantes n (%)	Evaluadores n (%)
Cirugía General	19 (82,6)	4 (33,3)
Medicina Familiar	14 (82,4)	8 (66,6)
Tocoginecología y ginecología	31 (86,1)	11 (78,6)
Total	64 (84,2)	23 (60,5)

Para la mayoría de los postulantes esta era la primera vez que participaban en entrevistas con esta metodología, el 26,5% (17/64) refirió haber tenido alguna experiencia de similar formato (16 en referencia al ECOE y 1 solo a MME).

En base al análisis de las respuestas de los postulantes al cuestionario se puede concluir que la mayoría consideró que la duración de las estaciones fue suficiente para demostrar sus ideas y fortalezas para el puesto, que las MME estuvieron libres de prejuicios culturales y de género y que la modalidad virtual no afectó el desempeño. Todos los evaluadores consideraron que el entrenamiento para su rol fue adecuado y la mayoría acordó que pudieron evaluar adecuadamente a los postulantes en el proceso de MMEv. El 17,2% de los postulantes y el 39% de los evaluadores consideró que la jornada de MME fue cansadora, cabe aclarar que los evaluadores estuvieron la jornada completa (dos circuitos, 4-5 hs aproximadamente) mientras que los postulantes solo en el circuito que les tocó participar (1.30h aproximadamente). El detalle de las respuestas se puede observar en las tablas 9 y 10.

Tabla 9. Respuestas de postulantes al cuestionario de aceptabilidad

Postulantes	Media ^a (DS)
La charla informativa previa acerca de la modalidad MME me resultó adecuada	4,75 (0,71)
Tuve tiempo suficiente para mostrar mis ideas y/o comportamientos en las estaciones	3,92 (0,96)
Las MME estuvieron libres de prejuicios culturales	4,30 (0,95)
Las MME estuvieron libres de prejuicios de género	4,38 (0,97)
Las MME me permitieron mostrar mis fortalezas para mi futuro desempeño como residente	4,02 (0,97)
El día de la entrevista no fue excesivamente cansador	3,84 (1,21)
Prefiero las MME sobre otros tipos de entrevistas como método de selección para el ingreso a la residencia	3,88 (1,09)
La modalidad virtual no afectó mi desempeño	4,27 (1,10)

^a Máximo puntaje de 5 (1: totalmente en desacuerdo a 5: totalmente de acuerdo).

Tabla 10. Respuestas de evaluadores al cuestionario de aceptabilidad

Evaluadores	Media ^a (DS)
La información que recibí sobre la estación me preparó adecuadamente como evaluador	4,65 (0,49)
La hoja de puntuación me permitió diferenciar entre los postulantes	4,09 (1,04)
Las MME permitieron evaluar las fortalezas de los postulantes para el futuro desempeño como residentes	4,13 (0,81)
Tuve tiempo suficiente para evaluar a los postulantes en mi estación	4,48 (0,85)
Incluir mayor número de evaluadores es una de las fortalezas de este método	4,74 (0,54)
El día de la entrevista no fue excesivamente cansador	3,30 (1,52)
La modalidad de MME virtuales permite realizar una buena valoración de los/as postulantes.	4,04 (0,93)

^a Máximo puntaje de 5 (1: totalmente en desacuerdo a 5: totalmente de acuerdo).

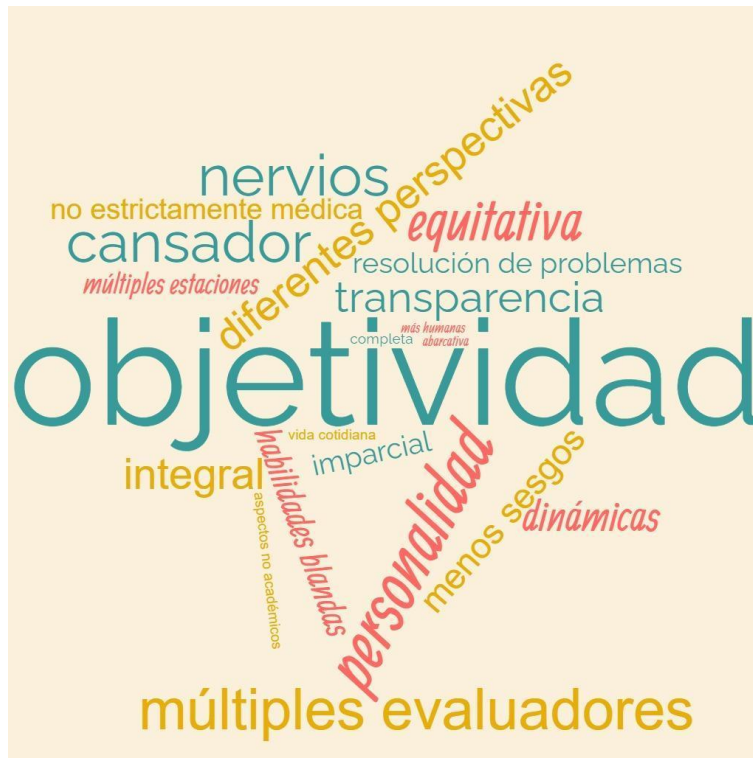
En cuanto a las ventajas mencionadas por los postulantes, se destacan principalmente la objetividad y posibilidad de ser evaluados en múltiples aspectos con esta metodología de entrevistas. Al mismo tiempo hacen referencia a considerarlo un método más justo y con menos posibilidad de sesgos (utilizaron términos como equitativa, más justa, transparencia, menos sesgos, evita acomodo, etc). En relación a las desventajas mencionan principalmente necesidad de más tiempo para desenvolverse en algunas estaciones, dificultades de la virtualidad y la imposibilidad de prepararse con anticipación para una entrevista con esta metodología. Algunos aspectos que para algunos fueron ventajas para otros no lo fueron relacionado a la duración y los nervios que les generaba. El detalle de los comentarios más relevantes se puede observar en la tabla 11 y en la figura 3.

Tabla 11. Ventajas y desventajas: selección de comentarios escritos por postulantes

Ventajas	Desventajas
Nos dio la posibilidad que todos seamos evaluados en los mismos aspectos, generando una forma más objetiva y equitativa .	Escaso tiempo , principalmente en las estaciones de entrevista tradicional. El tiempo es un factor a tener en cuenta
Disminuir la posibilidad de sesgos por relación entre el entrevistado y el entrevistador.	Pueden ser un poco largo y cansador (mencionado por 16 personas)

En caso de haberlos se diluyen en el número de estaciones.	
Me sacó la rigidez que las entrevistas me generaban	Pérdida de la impresión presencial del postulante
Realmente sentía nervios previo a la entrevista y que sea larga y en diferentes etapas, me permitió relajarme y poder desempeñarme mejor	Quizás la incertidumbre de tener una entrevista así por primera vez da bastantes nervios. Situación ansiógena.
Evalúa muchos aspectos de la personalidad de los postulantes que creo que serían claves para el desempeño como residentes.	Es difícil preparar una entrevista de este estilo y uno no sabe muy bien qué esperar o cómo debería responder.
Transparencia , brinda la posibilidad de conocer al postulante desde otras perspectivas	Desconocemos lo que están evaluando, te sentís inseguro, no llegas a demostrar tus capacidades.
Permite evaluar a una persona en situaciones cotidianas de la vida para poder determinar de esta manera como es la persona, sus características, atributos. Es una evaluación integral del postulante	La virtualidad dificulta un poco, por temas de audio entrecortado a veces, o el diley podía hacer que hable encima del entrevistador.
Se puede evaluar un lado más humano y más real de los postulantes . Los entrevistadores ven nuestras cualidades en la práctica en vez de simplemente escucharnos decir que tenemos cierta cualidad. Se evalúan actitudes humanas y no simplemente un curriculum.	No destaco ninguna desventaja (mencionado por 8 personas).
Me gusta que “sin querer queriendo” termina mostrándose uno como realmente es al ponerse en diversas situaciones cotidianas, sin la necesidad de que uno se describa.	Hubo pocas estaciones donde me permitieran hablar de mí.
Siento que el análisis para la selección del puesto es mucho más profundo que en una entrevista convencional que a veces es bastante superficial.	Desconocemos lo que están evaluando, te sentís inseguro, no llegas a demostrar tus capacidades.
MME nos dio la posibilidad que todos los postulantes seamos evaluados en los mismos aspectos, generando la forma más objetiva y equitativa para evaluarnos	
Anonimato	

Figura 3: Nube de palabras con los términos más frecuentes referidos por postulantes en sus comentarios escritos



A su vez, los evaluadores remarcan como principales ventajas la inclusión de múltiples entrevistadores con diferentes miradas y situaciones variadas a las cuales son expuestos los postulantes, posibilidad de reducir sesgos y realizar evaluaciones independientes. Como principal desventaja mencionan que les resulta cansador el proceso y que es una metodología que requiere gran organización. La tabla 12 ejemplifica los comentarios más representativos.

Tabla 12. Ventajas y desventajas: selección de comentarios escritos por evaluadores

Ventajas	Desventajas
<p>Homogeneidad en evaluación. Equidad.</p> <p>Todo el mundo pasa por situaciones casi idénticas. No te cansa como las entrevistas "convencionales".</p>	<p>La modalidad virtual desfavorece el encuentro.</p>
<p>Las múltiples miradas y múltiples situaciones seleccionadas, práctico al dividir la carga de trabajo para el entrevistador.</p>	<p>Tiempo y organización.</p> <p>Complejidad.</p>
<p>No conocer los nombres de los postulantes, evaluar otras capacidades no cognitivas.</p>	<p>Son muchos postulantes y mantener la atención por mucho tiempo se complica.</p>
<p>Permite observar características de</p>	<p>No sé me ocurren desventajas (referido por 7</p>

funcionamiento, individual y grupal, responsabilidad, compromiso, y la comprensión pertinente a las consignas de cada participante.	personas).
La cantidad de evaluadores . Esto permite varias observaciones y con ello minimiza errores . La objetividad de no conocer previamente a los postulantes.	Es un proceso muy largo y sería bueno intentar generar más pausas para la población de entrevistadores entre circuitos, o hacerlas en días diferentes.
Permite evaluar aspectos que muchas veces son difíciles de evaluar, permite minimizar las subjetividades al mantener el anonimato , permite detectar actitudes problemáticas y potencialmente peligrosas al evaluar también interacción entre pares.	No se si evalúa el perfil de residente que buscamos. Aun así creo que son mejores que las anteriores.
El número de observadores y que no se influyen entre ellos.	A veces el tiempo es insuficiente (de todos modos, el uso del tiempo también es un aspecto a tener en cuenta en la evaluación del candidato).

Discusión

El desarrollo del presente trabajo tuvo como objetivo principal evaluar la estructura factorial o de constructo de las MMEv. El AF evalúa la validez y busca resumir las interrelaciones existentes dentro del conjunto de variables del instrumento en una forma clara y segura. El diseño inicial de las MME contó con 5 variables (atributos no cognitivos): trabajo en equipo, razonamiento ético, argumentación, recepción de feedback y motivación hacia la especialidad e institución. A través del AFE se corroboró que los datos cargan en un solo factor (considerando, como suele hacerse, un valor mínimo de 1 eigenvalue para tener en cuenta a los factores) (Hair et al., 2018), por lo cual se considera una estructura unifactorial de la medición. Estos resultados son consistentes con los publicados por la Universidad de Ciencias de Malasia donde a partir del análisis factorial confirmatorio obtuvieron un constructo unidimensional (Abdul Rahim & Yusoff, 2016). Por otra parte, otra publicación sobre un análisis factorial sugiere un modelo de dos factores (comunicación oral y evaluación de problemas) altamente correlacionados (0,87), frente a estos resultados puede pensarse que son dos factores de un mismo constructo por la alta correlación, existiendo un constructo que carga en dos factores (Oliver et al., 2014). La idea de la medición de un constructo unidimensional no es exclusiva de las MME sino que también se encuentra para el

ECOE (Cunnington et al., 1996). Según ha indicado el Dr. Carlos Brailovsky, "estudios realizados en Canadá por el Royal College of Physicians and Surgeons y el Medical Council of Canada, el análisis factorial sobre resultados del ECOE a nivel nacional encontró también un modelo unifactorial" (comunicación personal, 18 de febrero de 2022).

A partir del análisis temático del contenido del corpus empírico (entrevistas y comentarios escritos) surgieron aspectos que contribuyen de manera similar a la valoración realizada por evaluadores: la primera impresión y la manera en la que resuelven la tarea. El fenómeno de la primera impresión (juicio realizado rápidamente y basado en mínima información) está bien estudiado en el ámbito de la psicología social y fue analizado recientemente en el escenario de las MME por Klusmann (2022), donde concluye que predice moderadamente medidas conductuales. También remarca que los evaluadores generalmente combinan su tarea específica descrita en las instrucciones de las MME con la pregunta autoimpuesta de si un candidato encajaría en el rol de médico (Klusmann et al., 2022). En nuestro trabajo surge esta primera impresión relacionada a aspectos de la personalidad (actitud, imagen, postura) y comunicación (verbal y no verbal: saludo, presentación, tono) del postulante aunque algunos evaluadores que participaron anteriormente en MME presenciales remarcaron que la virtualidad limitaba estas percepciones. Por otro lado, la resolución de la tarea está asociada a la manera en que comprende, analiza, jerarquiza y consecuentemente desarrolla la consigna planteada, estrechamente vinculada a la explicación e interacción con el otro en esa resolución.

Basándonos en el llamado efecto halo, es sabido que las diferentes puntuaciones (en este caso al atributo no cognitivo y al desempeño global) que un evaluador pueda asignarle a un postulante dentro de la estación están altamente correlacionadas ya que el evaluador desarrolla una impresión general y asignará puntajes a los diferentes ítems consistente con esa impresión general (Wood, 2014; Zaidi et al., 2016). A estos les resulta imposible de discriminar a pesar de que se realice el mejor entrenamiento, ya que al observar al postulante lo ven integralmente.

En esta oportunidad, y en comparación con las entrevistas de tipo panel, el peso de sesgos propios de los evaluadores, de antecedentes biográficos y académicos de los postulantes no surgieron como aspectos relevantes al momento de la definición de la puntuación. En todo caso, una de las grandes fortalezas de las MME es que los posibles sesgos individuales de los evaluadores se ven compensados al haber múltiples impresiones independientes (Pau et al., 2016).

La pregunta ¿qué miden las MME? tiene respuestas heterogéneas a partir de las publicaciones a lo largo de los años (Reiter & Eva, 2018) y posiblemente no tenga una única e inequívoca solución. Comúnmente, la intención es medir algún constructo de competencia personal. Frente a esto, Reiter y Eva (2018) señalan a las MME como un método de evaluación más que una herramienta, con un formato que puede generarse e implementarse en un número infinito de formas. Desde esta perspectiva, las MME deben pensarse como individualizadas para las necesidades de la institución. En nuestro caso, a través del AFE se obtuvo la medición de un constructo unifactorial y a través del análisis de contenido de los comentarios de los evaluadores, podemos decir que hacen foco principalmente en la primera impresión y la manera en la que resuelven la situación planteada en todos los aspectos.

Considerando los resultados obtenidos se constató que las MMEv resultaron un método confiable, válido y con buena aceptación por parte de postulantes y evaluadores de nuestro medio. Nuestros resultados coinciden con las revisiones sistemáticas sobre las características psicométricas de las MME realizadas en otros países hasta el momento (Pau et al., 2013; Rees et al., 2016; Yusoff, 2019). Para asegurar un ranking confiable, en el análisis de generalizabilidad la diferenciación entre postulantes debería ser de al menos del 15% en la faceta P (varianza de diferenciación) y el coeficiente G superior a 0,70. En nuestro caso el coeficiente G fue de 0,72 y la faceta P explicó el 16,7% de la varianza total.

La validez fue evaluada a partir de diferentes fuentes de evidencia empírica con respecto al contenido, proceso de respuesta y estructura interna. De esta manera obtuvimos diversos resultados que avalan su validez y están de acuerdo con otras publicaciones del tema (Abdul Rahim & Yusoff, 2016; Lemay et al., 2007; Pau et al., 2013).

Nuestros hallazgos en relación a la aceptabilidad de las MME virtuales son similares a lo publicado en la bibliografía, tanto postulantes como evaluadores las valoraron positivamente (Burgos et al., 2020; Hofmeister et al., 2008; Kok et al., 2021). Coincidieron en que les permitió demostrar y evaluar aptitudes para el rol de residente respectivamente, resaltaron la transparencia/objetividad y los postulantes remarcaron el hecho de evaluar aspectos no cognitivos como un punto importante dentro del perfil de un futuro médico.

Cabe destacar que algunas limitaciones del presente trabajo se relacionan con que es unicéntrico, se basa en la experiencia de las MMEv en tres residencias de una única institución, de gestión privada, de la Ciudad Autónoma de Buenos Aires lo que limitaría la generalización de los resultados. Al mismo tiempo, el posible conflicto de intereses existente si algún postulante y evaluador se conocen previamente (por haber rotado en el servicio, haberlo tenido de docente en el grado, etc.). En caso de que esto pasara consideramos que el contar con múltiples estaciones de evaluación independientes colabora a diluirlo, ya que por ser un hospital universitario con un ambiente de trabajo relativamente pequeño es imposible evitarlo completamente. Otra posible limitación en el análisis de la aceptabilidad es la tasa de respuesta de los evaluadores, la cual no es baja y permite el análisis y reporte de sus resultados de manera consistente con los reportados previamente en la bibliografía sin embargo no se puede dejar de mencionar que podría existir sesgo de selección, ya que contamos con la opinión de un segmento de los evaluadores y no del total de ellos.

Pese a lo mencionado, creemos que el principal aporte radica en la generación de evidencia para el nivel local, con gran validez interna. Asimismo, desde el punto de vista de la investigación cualitativa la validez está dada por diferentes aspectos como el compromiso en el trabajo de campo y la participación intensiva (Maxwell, 2012). Se utilizaron dos métodos de recolección de información de evaluadores (comentarios escritos y entrevistas) y de una muestra heterogénea en relación a experiencia, edad, especialidad. La información obtenida de las entrevistas fue relevada registrando la mirada de los distintos participantes. Las interpretaciones fueron realizadas respetando sus perspectivas sin alterar sus testimonios y el significado que ellos le atribuían. Consideramos que la presencia de

la investigadora principal afecta inevitablemente el medio observado y considerando que es docente y médica de la institución, intentamos reconocer esta incidencia e incorporarla (Maxwell, 2012). Para garantizar la validez y fiabilidad, teniendo en cuenta los posibles sesgos sobre el tema de investigación se recurrió a la codificación de la información obtenida de entrevistas y comentarios escritos por dos investigadoras independientes. En la investigación cualitativa se hace referencia a la "transferibilidad" en lugar de "generalizabilidad estadística" de las conclusiones refiriéndose a que la teoría desarrollada se puede extender a otros casos (Guba & Lincoln, 1989; Hammersley, 2018). En base a este criterio, consideramos que ciertos hallazgos de los resultados podrían ayudar a comprender y orientar problemáticas similares en los ámbitos de selección de residentes ya sea de manera ampliada en nuestra institución como en otras con sistemas de formación de especialistas médicos similares. Sin embargo, sabemos que no es posible realizar la extrapolación precisa de resultados.

Otra fortaleza de las MME (ya sean presencial o virtuales) es que son un método de entrevista estandarizado, con fuerte evidencia que lo respalda para la evaluación de aspectos no cognitivos a diferencia de lo que ocurre con las entrevistas tradicionales de tipo panel. En nuestro caso se logró la participación de tres residencias de la institución que demostraron interés en continuar utilizándolo y de esta manera poder involucrar a más médicos del servicio en la selección de sus futuros residentes y colegas. Otra ventaja está vinculada al rol de evaluación "ciego" (no contar con el nombre, antecedentes o notas académicas de los postulantes) e independiente en cada una de las estaciones. Por último las publicaciones sobre análisis factorial que existen al momento son todas sobre MME realizadas para ingresantes a escuelas de Medicina (grado), no se ha realizado sobre MME virtuales en la selección de residentes (posgrado).

Nuevos estudios serán necesarios para explorar si los médicos seleccionados por este método tienen un buen desempeño en el transcurso del programa de formación. También sería posible evaluar específicamente el fenómeno de la primera impresión para la puntuación en nuestro ámbito de trabajo. De especial interés, sería explorar el costo-beneficio de este método comparado con los otros métodos de selección de residentes.

Conclusiones

Esta tesis permite concluir que las MMEv miden un constructo unidimensional. Desde el punto de vista de los evaluadores los dos principales aspectos que influyen la evaluación son la primera impresión y la resolución de la situación planteada en la estación. Se reafirmó como un método confiable, válido y bien aceptado, que permite la generación de múltiples impresiones individuales de los postulantes en aspectos no académicos necesarios para el rol de médico especialista en un relativamente corto período de tiempo.

Anexo I: Tabla especificaciones (blueprint)

	Estación	Atributos no cognitivos					Desempeño global	Total
		Motivación	Argumentación	Trabajo en equipo	Ética	Recepción de feedback		
1	Motivación por la especialidad	8					2	10
2	Actividades extras	8					2	10
3	Noticia de actualidad		8				2	10
4	Dupla: dibuja		8				2	10
5	Tarea con ayudante			8			2	10
6	Rol de líder			8			2	10
7	Negociación			8			2	10
8	Redes sociales				8		2	10
9	Responsabilidad médica				8		2	10
10	Error				8		2	10
11	Historieta					8	2	10
12	Dupla: Guía					8	2	10
	Puntaje total por atributo	16	16	24	24	16	24	120
	Ponderación %	13.3	13.3	20.0	20.0	13.3	20.0	100,0%
	Ponderación objetivo	15% ± 5%	15% ± 5%	20% ± 5%	20% ± 5%	15% ± 5%	15% ± 5%	
13	Terapia alternativa		8				2	10
14	Consumo				8		2	10
	Puntaje total por atributo	16	24	24	32	16	28	140
	Ponderación %	11.4	17.1	17.1	22.9	11.4	20.0	100,0%

Las estaciones 13 y 14 solo fueron utilizadas para Tocoginecología y Ginecología Quirúrgica por tener mayor cantidad de postulantes a entrevistar. La ponderación de cada dominio evaluado al agregar estas 2 estaciones extras se mantuvo en el rango pre establecido (ponderación objetivo) en el diseño de la tabla de especificaciones.

Anexo II: Ejemplo de estaciones y ficha de evaluación

***ESTACIÓN 1**

Nombre de la estación: Motivación hacia la especialidad

Dominio que evalúa: Motivación hacia la práctica de la Medicina Familiar

Tipo de estación: uno a uno

Duración: 7 minutos

Escenario: Se realizará una entrevista semiestructurada con las siguientes preguntas disparadoras:

-Tiene 2 minutos para presentarse libremente.

-Luego debe contestar a las siguientes preguntas que le hará el entrevistador/a:

¿Por qué elegiste la especialidad ...?

¿Cuáles son las características distintivas de la especialidad que motivaron tu elección?

¿Cuáles son los motivos por los cuales te interesaría hacer la residencia con nosotros?

Consigna para el/la entrevistador/a

Esta estación está diseñada para evaluar: Motivación hacia la especialidad y desempeño global

-Al inicio de la entrevista usted debe postear en el chat (como mensaje privado al participante) la siguiente consigna:

-Tiene 2 minutos para presentarse libremente.

-Luego debe contestar a las siguientes preguntas que le hará el entrevistador/a:

¿Por qué elegiste la especialidad ...?

¿Cuáles son las características distintivas de la especialidad que motivaron tu elección?

¿Cuáles son los motivos por los cuales te interesaría hacer la residencia con nosotros?

Su ROL:

-Ni bien ingresa el/la postulante a la sala de zoom debe pegar en el chat las instrucciones.

El objetivo es escuchar la auto presentación del postulante en 2 minutos y luego hacer las siguientes preguntas, pidiendo al postulante que expongan su respuesta a cada una.

-¿Por qué elegiste ...?

-¿Cuáles son las características distintivas de la especialidad que motivaron tu elección?

-¿Cuáles son los motivos por los cuales te interesaría hacer la residencia con nosotros?

Opciones de preguntas para ampliar si tiene tiempo:

*Además de lo asistencial hay **otro área que te interese**, por ej. investigación, docencia, salud comunitaria, salud pública, etc.*

Puntuación: Usted debe proporcionar dos puntuaciones: motivación y desempeño global.

Motivación hacia la especialidad:

Postulante Excelente: Conoce la especialidad. Se informó acerca de los alcances de la especialidad. Tuvo algún tipo de contacto con la práctica de la especialidad. Conoce diferentes servicios y ámbitos donde se desarrolla la especialidad. Puede diferenciarla de otras especialidades como clínica.

Postulante Pobre: No conoce la especialidad ni en lo que trabaja un médico de familia. No tuvo ningún contacto con la especialidad. No es claro cuando explica por qué elige esta y no otra especialidad.

Desempeño Global:

Aquí se trata de evaluar el desempeño del postulante en forma global, teniendo en cuenta todos los aspectos del caso

Postulante excelente: demuestra juicio, síntesis, claridad, coherencia, conocimiento de la especialidad y de la institución.

Postulante Pobre: quien no logra demostrar alguno de los aspectos previamente descriptos.

Banderas rojas de la estación:

-Ve a la especialidad como un puente para llegar a otra especialidad.

-Este es el único hospital en el que rindió para esta especialidad.

***ESTACIÓN 4**

Nombre de la estación: Redes sociales

Dominio que evalúa: Razonamiento ético

Tipo de estación: uno a uno

Duración: 7 min

Escenario: Ud. está en una cena con amigos/as y le muestran un video donde se ve a un compañero de otra residencia claramente borracho, con muy mal aspecto y vomitando a la salida de un boliche. Ud. conoce a esta persona con quien ha tenido varios problemas durante las guardias y se lo cruza habitualmente en el hospital. Sus amigos están por subir el video a Tiktok como algo anecdótico y divertido, sabiendo que lo verán muchas personas del trabajo y la pareja del protagonista.

1) Considere cuáles serían las situaciones problemáticas para este caso.

2) Plantee posibles alternativas de manejo a estas situaciones.

Consigna para el/la evaluador/a

Esta estación está diseñada para evaluar: razonamiento ético y desempeño global.

-Al inicio de la entrevista usted debe postear en el chat (como mensaje privado al participante) la siguiente consigna:

Ud. está en una cena con amigos/as y le muestran un video donde se ve a un compañero de otra residencia claramente borracho, con muy mal aspecto y vomitando a la salida de un boliche. Ud. conoce a esta persona con quien ha tenido varios problemas durante las guardias y se lo cruza habitualmente en el hospital. Sus amigos están por subir el video a Tiktok como algo anecdótico y divertido, sabiendo que lo verán muchas personas del trabajo y la pareja del protagonista.

1) Considere cuáles serían las situaciones problemáticas para este caso.

2) Plantee posibles alternativas de manejo a estas situaciones.

Su ROL: El objetivo es valorar las opiniones de los candidatos sobre este tema en 7 minutos.

- Usted debe hacer las siguientes preguntas, pidiendo al candidato que expongan su respuesta a cada una.
 1. ¿Qué situación/nes problemáticas identificas en este caso?
 2. ¿Cómo actuarías en esta situación?
 3. ¿Consideras que es apropiado subir fotos/videos a las redes sociales donde aparecen otras personas? ¿Pensas igual si en el video la persona está en mal estado (ej. borracho) que si es en una situación más "normal" donde no hay conflicto aparente? Por qué?

4. ¿Habrías con tus amigos/as acerca del posteo? Si esto te pone en una situación contra tu grupo de amigos/as, actuarías igual?
5. ¿Lo impedirías? ¿Si no podés hacerlo y sabes que se sube a Tiktok le avisas al implicado? ¿Cómo, Por qué?
6. En este caso es una persona a quien conoces, pero si fuera un video igual de un desconocido, ¿cómo actuarías?
7. ¿Qué pensas de la privacidad, del uso de la imagen de otros?
8. ¿Cómo consideras que debería ser un uso responsable de redes sociales pensando en esta situación?

Puntuación: Usted debe proporcionar dos puntuaciones: razonamiento ético y desempeño global.

Razonamiento ético

Candidato Excelente: Expresa una comprensión equilibrada de la situación problemática planteada en el caso. Expresa una comprensión equilibrada de las consecuencias de cada una de las posibles soluciones. Valora las consecuencias de cada decisión y los costos para cada uno de los involucrados. Plantea estrategias alternativas apropiadas. Plantea adecuadamente las razones para la elección de cada una de las estrategias alternativas

Candidato Pobre: mala comprensión de la situación problemática planteada en el caso. No le es posible exponer las consecuencias de cada una de las posibles soluciones. No le es posible sugerir estrategias alternativas

Desempeño Global:

Aquí se trata de evaluar el desempeño del candidato en forma global, teniendo en cuenta todos los aspectos del caso

Candidato excelente demuestra juicio, síntesis, efectividad y eficiencia.

Candidato Pobre: es aquel que no logra demostrar alguno de los aspectos previamente descriptos.

Banderas rojas de la estación:

-No interviene, hace como que no vio nada, no es su problema.

***ESTACIÓN 5**

Nombre de la estación: Historieta

Dominio que evalúa: Capacidad para recibir feedback

Tipo de estación : uno a uno

Duración: 7 minutos

Escenario: El participante mirará una historieta y luego deberá explicar y argumentar que entendió, que mensaje le transmitió. Luego el entrevistador le dará feedback.

Consigna para el/la entrevistador/a

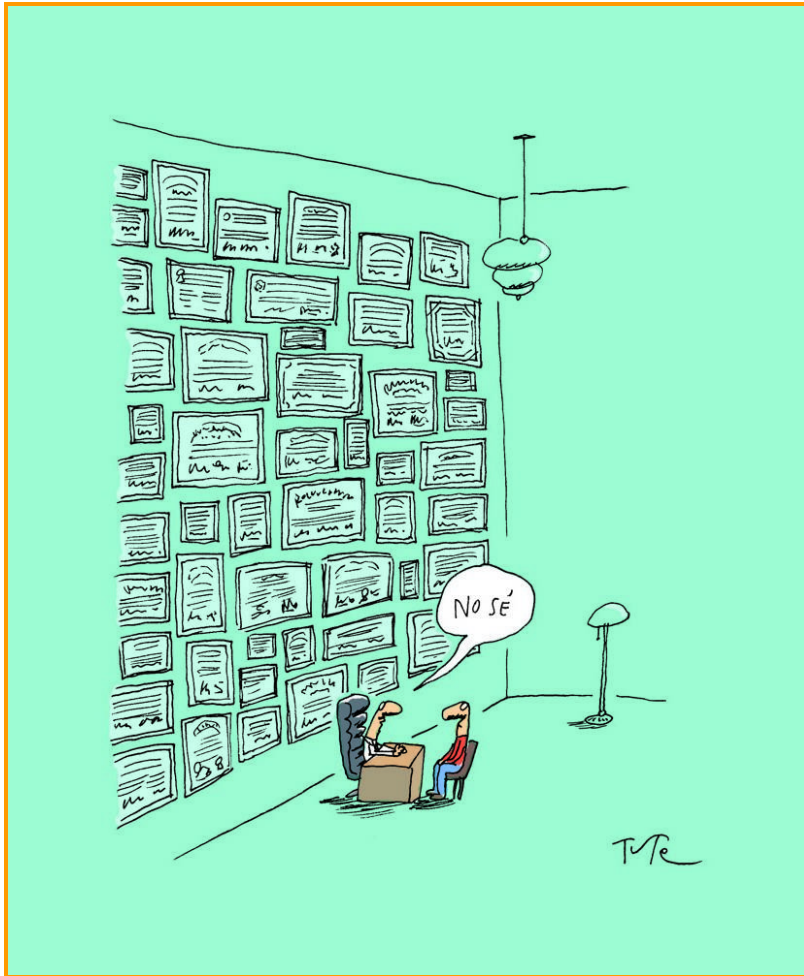
Esta estación está diseñada para evaluar: Recepción de feedback y desempeño global

-Al inicio de la entrevista usted deberá compartir la pantalla mostrando al postulante la siguiente consigna e historieta que le daremos en una diapositiva de ppt:

Mire la historieta que el/la entrevistador/a le va a compartir y luego exponga en 3 minutos, en relación a las siguientes preguntas:

- *¿Qué es lo primero que se le viene a la mente al verla?*
- *¿Qué aspectos relevantes encuentra en la historieta (puede ser sobre dibujos o diálogo)?*
- *¿Sobre qué temas puede reflexionar luego de mirarla?*
- *¿Qué mensaje considera que deja esta historieta?*

Al finalizar el/la entrevistador/a le dará una devolución en relación a sus comentarios.



Su ROL:

El objetivo es **valorar cómo responde y argumenta al feedback.**

1. Usted debe permitir que el/la postulante comience a hablar, si ve que se queda callado puede estimularlo leyendo las preguntas de la consigna.
2. Luego de escuchar al postulante por aprox. 3 minutos debe darle Feedback sobre su exposición.

INDICACIONES PARA EL FEEDBACK:

Para intentar que todos/as los/as postulantes reciban una devolución similar, se sugiere realizarlo de la siguiente forma, tiene un minuto para hacerlo:

- Deberá decirle al postulante cuando se cumplan los 3 minutos: "Ya finalizó el tiempo asignado para esta tarea, ahora te voy a hacer algunos comentarios acerca de tu desempeño".
- Iniciar con Feedback positivo: Nombrar un solo aspecto, ser objetivo en relación a las conductas que tuvo el postulante. Ej.: "Tu exposición fue clara con respecto a aspecto" "Fuiste imaginativo al mencionar"
- Luego Feedback negativo: Nombrar un solo aspecto a mejorar. Ej. "Lo que deberías mejorar sería..." "Te faltó dar alguna apreciación sobre... tal aspecto de la historieta" "Tuviste pocos argumentos al momento de explicar solo mencionaste tal y tal cosa..." "Te repetiste en tal aspecto..."

0. Al finalizar el feedback **debe preguntarle al postulante que opina acerca de este feedback y dar la oportunidad a que responda a esto.**

- Ej. Podría decirle ¿cómo te autoevalúas en esta consigna? ¿Coincidís con lo que te acabo de decir? ¿Qué crees podrías haber hecho mejor?

Puntuación: Usted debe proporcionar dos puntuaciones: Recepción de feedback y desempeño global.

No debe juzgar el conocimiento, ni las opiniones sino las conductas específicas relacionadas con los dominios a evaluar.

Capacidad para recibir feedback

Postulante Excelente: Escucha los argumentos del entrevistador cuando le da feedback. Reacciona adecuadamente al feedback recibido. Es receptivo a las críticas. Es capaz de realizar una autoevaluación adecuada, es reflexivo.

Postulante Pobre: No escucha los argumentos del entrevistador cuando le da feedback. Reacciona inadecuadamente al feedback recibido. Es poco receptivo y muy reactivo a las críticas. No es capaz de realizar una autoevaluación adecuada.

Desempeño Global:

Aquí se trata de evaluar el desempeño del postulante en forma global, teniendo en cuenta todos los aspectos del caso

Postulante Excelente: demuestra juicio, síntesis, claridad, coherencia y eficiencia.

Postulante Pobre: es aquel que no logra demostrar alguno de los aspectos previamente descriptos.

Banderas rojas de la estación:

-El candidato reacciona al feedback de manera agresiva o desproporcionada para la situación.

***ESTACIÓN 6**

Nombre de la estación: Tarea con ayudante: Ejercicio con ÁBACO

Dominio que evalúa: Trabajo en equipo/Habilidades interpersonales

Tipo de estación: Interactiva con actriz

Duración: 7 min

Escenario: Ud. es voluntario/a en un centro de día al que asisten adolescentes con dificultades en el aprendizaje. En esta estación Aldana, una de las jóvenes que asisten al centro, deberá realizar junto a Ud. la siguiente actividad con un ábaco.

Consigna para la actriz

En esta estación se le asigna al postulante una tarea para que realice junto a Ud. Ud. estará conectada en su domicilio con una computadora a través de zoom.

Ud. será una joven con dificultades en el aprendizaje que asiste a un centro de día, tiene que realizar una actividad con un ábaco (lo tendrá en su domicilio) y el postulante tiene que ayudarla.

Tiene que colocar la cámara de tal manera que se vea bien su cara y el ábaco para mostrar los movimientos que vaya realizando.

Rol de la actriz

Será Aldana, la joven (con dificultades del aprendizaje).

Las estaciones se repetirán de igual manera con cada postulante, con cada uno de ellos Ud. piense que es la primera vez que realiza esta actividad, *pídales que le lean despacio las instrucciones y se confundirá en algunos pasos.*

Cualquier cosa que el postulante le pida que haga, hágalo. Una parte importante de su rol es no ser "tan útil" y expresar cierta confusión frente a las tareas asignadas.

1. Pídale al postulante que le lea las instrucciones y una vez que termine empiece a sacar y poner en cualquier orden las piezas.

0. **Episodios de confusión:**

Deberá expresar confusión en 2 oportunidades, son instancias para que el/la postulante le explique la actividad o parte de ella.

Ejemplos:

- Cuando le pidan que separe las piezas por color mezcle algunas y si el postulante le explica como diferenciar por colores, puede mostrar una pieza roja y preguntar ¿esta es azul? Confundiendo el color.

- Si el postulante le dice que ordene las piezas en el ábaco dígame que no entiende en que palito debe ponerlas.

-Si le menciona derecha e izquierda, puede confundir cual es cual.

El postulante debería explicarle amablemente.

En uno de los episodios de confusión: la primera vez que le explique diga que no entiende y haga que vuelva a explicarlo, cuando se lo explique por segunda vez Ud. diga: "Perdón, ahora entendí".

0. **Episodios de errores:**

Deberá equivocarse 2 veces. Deberán parecer errores inocentes, si el postulante los descubre Ud. deberá actuar como avergonzada y compungida (Uh, mil perdones, estoy haciendo todo mal... Estoy haciéndote perder el tiempo, perdón...). Es importante que tenga los dos comportamientos: equivocarse 2 veces y estar compungida por los errores y pedir perdón varias veces. Si el postulante no nota los errores continuar con la actividad.

Ejemplos de errores:

-Mezclar los colores al colocarlos en el ábaco.

-Colocar más de una ficha del color indicado, ej: si le dice por ej que coloque 3 amarillas en un palito, Ud. coloque 4.

Tiene que aparentar con muchas ganas de ayudar al postulante.

Qué hacer si...

*El postulante se equivoca con las indicaciones que le está dando (lee mal las instrucciones o se confunde) sígale la corriente.

*El postulante se enoja con Ud. por sus errores. Sígale la corriente, si en algún momento lo trata mal el evaluador podrá intervenir.

*El postulante pierde la paciencia al explicarle y que Ud. se equivoque varias veces muéstrese compungida.

Consigna para el observador/a

En esta estación se le asigna al postulante una tarea para que realice en poco tiempo que consiste en ordenar las piezas de un ábaco con la ayuda de un compañero (con ciertas dificultades del aprendizaje).

Esta estación está diseñada para evaluar: Trabajo en equipo/Habilidades interpersonales y desempeño global

-Al inicio de la entrevista usted debe postear en el chat (como mensaje privado al participante) la siguiente consigna:

Ud. es voluntario en un centro de día al que asisten jóvenes con dificultades en el aprendizaje. En este caso Aldana, una de las jóvenes, le pide ayuda para realizar una actividad con un ábaco. Deberá darle indicaciones para que ella pueda ordenar las piezas siguiendo las instrucciones:

1) *Separar todas las piezas por color.*

2) *Ordenar las piezas en los tres palitos (de derecha a izquierda de Aldana):*

-3 azules y 1 amarilla

-2 azules y 3 rojas

-5 rojas y 2 amarillas

Tiene 7 minutos para realizar esta tarea.

Su ROL:

Ud. deberá evaluar la interacción entre el postulante y la actriz. Se le adjuntó previamente el libreto del actor para que pueda familiarizarse con las consignas que le hemos dado.

Puntuación:

Usted debe proporcionar dos puntuaciones: trabajo en equipo y desempeño global.

- *¿Cuál sería una buena estrategia de inicio?* Presentarse con la actriz, leer las instrucciones de manera completa en voz alta para que la joven también las escuche, pedirle que le muestre el ábaco si no se ve bien, repreguntar si comprendió las indicaciones, etc. La actriz va a haber practicado previamente (el postulante no lo sabrá).

Descripción de los dominios:

Trabajo en equipo/Habilidades interpersonales:

Postulante Excelente: Cómodo con el rol de líder. Le brinda instrucciones claras y lógicas a la actriz. Adecuado trato cuando la actriz se equivoca. Trabaja con calma y efectivamente bajo presión. Entiende el trabajo en equipo. Le da espacio a la actriz para la interacción.

Postulante Pobre: Discomfort con el rol de líder. Falta de confianza y claridad al brindar las instrucciones a la actriz. Trato inadecuado cuando la actriz se equivoca. No puede trabajar bajo presión. Nerviosismo excesivo frente a errores del otro. No plantea la actividad como un trabajo en equipo.

Desempeño Global: Aquí se trata de evaluar el desempeño del postulante en forma global, teniendo en cuenta todos los aspectos del caso.

Postulante excelente: demuestra juicio, síntesis, claridad, efectividad y eficiencia.

Postulante Pobre: es aquel que no logra demostrar alguno de los aspectos previamente descriptos.

Banderas rojas de la estación:

-El postulante trata de mala manera a la actriz.

-El postulante desiste de ayudarla con la actividad.

***ESTACION 11**

Nombre de la estación: Dibujo en duplas: POSTULANTE QUE DIBUJA

Dominios que evalúa: Argumentación

Tipo de estación: interactiva, doble

Duración: 7 min

Escenario: En esta estación un postulante deberá guiar con indicaciones verbales a otro para que realice un dibujo en 3 minutos, al finalizar quien dibujó deberá darle una devolución de su desempeño como guía y discutir dificultades surgidas durante la tarea.

Consigna para el observador/a

En esta estación interactúan dos postulantes a la residencia, uno da indicaciones para realizar un dibujo y el otro es el dibujante que luego le dará feedback al que lo guío.

Esta estación está diseñada para evaluar: Argumentación y desempeño global

-Al inicio de la entrevista usted debe postear en el chat (como mensaje privado al participante) la siguiente consigna:

En esta estación **UD. SERÁ GUIADO POR OTRO/A POSTULANTE PARA REALIZAR UN DIBUJO** (use lápiz y papel).

Durante 3 minutos lo guiará verbalmente para dibujar un elemento (su colega tiene prohibido utilizar algunas palabras) y Ud. puede realizar preguntas para que responda solo por Si o NO. Al finalizar el tiempo el evaluador/a les avisará, muestre a la cámara como quedó el dibujo.

Luego UD. TENDRÁ 2 MINUTOS PARA DARLE UNA DEVOLUCIÓN A SU COLEGA EN RELACIÓN A CÓMO FUE EL DESEMPEÑO DANDO INSTRUCCIONES (Ej. qué indicaciones le fueron útiles y que aspectos podría mejorar) y discutir entre ambos dificultades que puedan haber observado.

AL FINALIZAR LA ESTACIÓN PERMANECERÁ ya que tomará el rol de guía y repetirá la estación con otro/a colega.

*Le mostramos como evaluador/a un ej. del estilo de dibujo que deberá realizar y las indicaciones que recibe el otro postulante, el día del entrenamiento haremos una práctica.

1- Debe dar indicaciones para que su colega dibuje un **RELOJ** similar al siguiente:



2- **NO puede mencionar** las siguientes palabras:

HORA
TIEMPO
DESPERTADOR
AGUJAS
MUÑECA

3- Solo puede responder a las dudas de su colega con **SI o NO**.

Su ROL como evaluador:

Ud. debe evaluar solo al que dibuja en relación a cómo le da una devolución del desempeño al colega que lo guío.

Otro evaluador será el encargado de puntuar al postulante que guía y recibe el feedback.

Al finalizar los 3 minutos Ud. debe avisarle y pídale al dibujante que le dé una devolución en relación a cómo fue hecha la guía y pídale que discutan acerca de la interacción que tuvieron.

*¿Cuál sería una buena estrategia para ser guiado y luego dar una devolución a su compañero?

- Escucha atentamente la guía de su compañero
- Pregunta si no comprende algo
- Menciona puntos positivos en relación a la guía que recibió
- Menciona puntos a mejorar en relación a la guía que recibió de manera clara, respetuosa.
- Escucha la reacción al feedback que brindó.

Puntuación:

Usted debe proporcionar dos puntuaciones: argumentación y desempeño global.

No debe juzgar el conocimiento, ni las opiniones sino las conductas específicas relacionadas con los dominios a evaluar.

Argumentación:

Postulante Excelente: Excelente enfoque lógico. Claro análisis de las problemáticas que surgen de la situación. Comunica de manera adecuada sus pensamientos y argumentos al evaluador. Plantea diferentes posibilidades de análisis. Plantea adecuadamente costos-beneficios de la problemática planteada. Tono apropiado. Mantiene el contacto visual. Es profesional y comprometido. Presenta los argumentos de forma clara, lógica y con confianza.

Postulante Pobre: Pobre enfoque lógico. Presenta los argumentos de manera desorganizada. No justifica los argumentos. Pocos o ningún análisis de planes posibles. No plantea costos beneficios de cada alternativa. Tono inapropiado. No mantiene contacto visual. Es poco profesional y comprometido. No presenta los argumentos de forma clara, lógica y con confianza.

Desempeño Global:

Aquí se trata de evaluar el desempeño del candidato en forma global, teniendo en cuenta todos los aspectos del caso.

Candidato excelente: demuestra seguridad, juicio, síntesis, efectividad y eficiencia.

Candidato Pobre: es aquel que no logra demostrar alguno de los aspectos previamente descriptos.

Banderas rojas de la estación:

-El candidato trata de mala manera al colega.

-Permanece callado sin posibilidad de dar ninguna devolución.

***Ejemplo de Ficha de evaluación**

Número del postulante:

Evaluador/a:

Marcar con una cruz en las siguientes escalas donde considera que se encuentra el desempeño del postulante para cada dominio.

Trabajo en equipo:									
1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pobre								Excelente	
Desempeño Global:									
1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pobre								Excelente	
Banderas rojas: SI									

¿Cómo construyó la valoración del desempeño?

Le solicitamos que explique con sus palabras tratando de identificar cuáles fueron los elementos en los que se basó para puntuar a el/la postulante.

Anexo III: Guía de entrevista

Introducción:

Mi nombre es Ma. de las Nieves Ganiele, soy la investigadora principal del proyecto "Múltiples Mini-Entrevistas virtuales como método de selección de residentes", el cual está enmarcado en mi tesis de la maestría en educación para profesionales de la salud. Estamos realizando este estudio para explorar qué atributos evalúan los/as entrevistadores/as para asignar un valor al desempeño de los postulantes en las MME virtuales.

El protocolo de este trabajo fue aprobado por el Comité de Ética de Protocolos de Investigación Universitaria (CEPIU), bajo el nro. 0016-22.

Toda la información que obtengamos será manejada en forma confidencial, siguiendo normas locales de confidencialidad de los datos obtenidos y su alcance con el cumplimiento de la ley 25.326.

Participar no conlleva ningún riesgo y los investigadores no obtendremos ningún beneficio económico del mismo. Probablemente se publicarán los resultados de este estudio, pero en ningún caso se utilizarán los nombres de los participantes, sólo las características de los grupos participantes. El único uso de la información será proporcionar datos para el estudio académico.

Experiencia general en las MME

1. ¿Cómo evaluó de manera general la experiencia de las MME virtuales?
2. ¿Qué ventajas encuentra?
3. ¿Qué desventajas?
4. ¿Qué piensa de que se hagan de manera virtual? *Hacerla si participaron en MMEp y MMEv, que comparen las dos modalidades con pros y contras.*

Foco en la entrevista en la cual participó

Pensando en la estación en la cual participaste, que apuntaba a evaluar *ARGUMENTACIÓN* y desempeño global:

5. Realmente ¿consideras que la estación te permitió diferenciar entre ARGUMENTACIÓN y DESEMPEÑO GLOBAL? ¿Crees que pudiste separarlos/diferenciarlos al momento de puntuar? Cómo construiste el puntaje final para ARGUM..?
6. ¿Cómo construyó la valoración del DESEMPEÑO global del postulante?
7. ¿Qué buscaba en el o la postulante? *Relacionado con el perfil del residente*

- a. ¿Cuáles fueron los elementos en los que se basó para puntuarlo?
 - b. ¿Qué ponderó como relevante?
 - c. ¿Consideras que el diseño de la estación influyó en lo que terminaste buscando en el postulante? ¿Te permitió evaluar lo que vos considerabas importante? ¿Mirabas otra cosa que no estuviera especificada en la estación? *Para ver si el evaluador buscaba otro atributo que no era el propio de la estación*
8. Hubo algún factor negativo (actitud, verbal, etc) que pueda invalidar al resto
 9. ¿Tuvo en cuenta algún componente del lenguaje no verbal, de qué manera?
 10. ¿Cómo se maneja al momento de calibrar la selección del puntaje?
Ej. vio a los primeros 3 y después ya podía diferenciar a los mejores de los peores? Al final cambió las puntuaciones una vez que vio a todos?
 11. Hay alguna situación que quieras contarme como para ejemplificar que piensas/observas/valoras al momento de evaluar.
 12. ¿Qué sesgos reconoces en vos mismo que traes a las entrevistas de selección y son difíciles de anular?
(mandame un audio x WA si se te ocurren más ideas al reflexionar)

Anexo IV: Modelo de cuestionarios de aceptabilidad

El cuestionario se enviará a todos los participantes vía mail una vez finalizado el proceso de entrevistas. A continuación se encuentra el mensaje que irá en el cuerpo del mail y el cuestionario propiamente dicho que se volcará en un formulario de Google Drive que no recabará información que permita reconocer a quien la responda para mantener el anonimato. Los datos obtenidos serán analizados una vez que el ranking final de ingreso se haya publicado para asegurar que no interfiere en la selección.

Múltiples Mini Entrevistas (MME) en el ingreso a residencias en el Hospital Italiano de Buenos Aires

Estimados/as, en el marco del proyecto "Múltiples Mini-Entrevistas virtuales como método de selección de residentes", a cargo de Ma. de las Nieves Ganiele, investigadora principal, se está realizando una investigación en relación a parámetros psicométricos y aceptabilidad de este método. Este sistema de entrevistas es utilizado ampliamente en otros países y está respaldado por evidencia publicada en el área de la educación médica.

Con este fin le pedimos su ayuda completando un breve cuestionario que les llegará a su email. En la misma se les solicita marcar el grado de acuerdo que considere tener para cada enunciado dentro de la escala propuesta de 5 ítems.

El cuestionario es anónimo, confidencial y los datos obtenidos serán utilizados exclusivamente con fines de investigación. La colaboración en el llenado del cuestionario no afecta el resultado final del proceso de selección de residentes, podrá retirar su consentimiento en cualquier momento.

Si está dispuesto a contestar puede hacerlo a continuación, accediendo a través del link que encontrará debajo.

Su experiencia es de suma ayuda para nuestro trabajo.

Muchas Gracias!

Por dudas o consultas comunicarse con: Ma. de las Nieves Ganiele
maria.ganiele@hospitalitaliano.org.ar proyecto aprobado por el Comité de Ética de Protocolos de Investigación Universitario (CEPIU) con número de aprobación: 0016-22.

Para postulantes

- Edad: _____
- Especialidad para la que se postula:
 - cirugía general
 - ginecología quirúrgica
 - tocoginecología
 - medicina familiar
- Género:
 - Mujer-cis
 - Mujer-trans
 - Varón-cis
 - Varón-trans
 - No binario
 - Prefiero no decirlo
 - Otro
- Universidad de la cual egresó: _____
- Marque con una cruz su grado de acuerdo con los siguientes enunciados

MME: Múltiples Mini Entrevistas

	Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo
La charla informativa previa acerca de la modalidad MME me resultó adecuada					
Tuve tiempo suficiente para presentar mis ideas en las estaciones					
Las MME están libres de prejuicios culturales					

Las MME están libres de prejuicios de género					
Las MME me permitieron presentar mis fortalezas para mi futuro desempeño como residente					
El proceso de entrevistas NO fue excesivamente cansador					
Prefiero las MME sobre otros tipos de entrevistas como método de selección para el ingreso a la residencia					
La modalidad virtual no afectó mi desempeño.					

- ¿Participó previamente de una sistema de evaluación similar (MME, ECOE, OSCE)?
- Si, ECOE
- Si, MME
- No
- Mencione ventajas de las MME

- Mencione desventajas de las MME

- Si tuviera que elegir, ¿qué modalidad de entrevistas prefiere ?
VIRTUAL/PRESENCIAL/INDISTINTO
- ¿Tuvo dificultades técnicas por la modalidad virtual? SI/NO
- Si tuvo dificultades: obtuvo el apoyo necesario del equipo organizador para su resolución? SI/NO/NS NC

- Si tiene alguna sugerencia para mejorar el proceso de MME puede escribirla aquí: _____

Para evaluadores

- Edad: _____
- Especialidad para la cual entrevistó:
 - cirugía general
 - ginecología quirúrgica
 - tocoginecología
 - medicina familiar
- Género:
 - Mujer-cis
 - Mujer-trans
 - Varón-cis
 - Varón-trans
 - No binarie
 - Prefiero no decirlo
 - Otro
- ¿Participó como entrevistador/a previamente de las Múltiples Mini Entrevistas (MME)?
 - Si, solo presencial
 - Si, solo virtual
 - Si, presencial y virtual
 - Es la primera vez
- Marque con una cruz su grado de acuerdo con los siguientes enunciados

	Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo
La información que recibí sobre la estación me preparó adecuadamente como evaluador/a					

La hoja de puntuación me permitió diferenciar entre los postulantes					
Las MME permiten evaluar las fortalezas de los postulantes para el futuro desempeño como residente					
Tuve tiempo suficiente para evaluar a los postulantes en mi estación					
Incluir mayor número de evaluadores es una de las fortalezas de este método					
El día de la entrevista NO fue excesivamente cansador					
La modalidad de MME virtuales permite realizar una buena valoración de los/as postulantes.					

- Mencione ventajas de las MME

- Menciones desventajas de las MME

- ¿Tuvo dificultades técnicas por la modalidad virtual? SI/NO
- Si tuvo dificultades: obtuvo el apoyo necesario del equipo organizador para su resolución? SI/NO/NS NC
- Si tiene alguna sugerencia para mejorar el proceso de MME puede escribirla aquí: _____

Anexo V: Modelo de invitación y consentimiento oral para las entrevistas a evaluadores/a

Mi nombre es Ma. de las Nieves Ganiele, soy la investigadora principal del

proyecto " Múltiples Mini-Entrevistas virtuales como método de selección de residentes"; el cual está enmarcado en mi tesis de la maestría en educación para profesionales de la salud. Estamos realizando este estudio para explorar qué atributos evalúan los/as evaluadores/as para asignar un valor al desempeño de los postulantes en las MME virtuales. El protocolo de este trabajo fue aprobado por el Comité de Ética de Protocolos de Investigación Universitaria (CEPIU), bajo el nro. 0016-22.

Estoy realizando entrevistas individuales con una duración aproximada 45-60 minutos, que podrán ser presenciales o virtuales, con fecha y hora a definir de manera consensuada entre ambas partes. Las mismas serán audio-grabadas para su posterior análisis, previo consentimiento oral informado.

Toda la información que obtengamos será manejada en forma confidencial, siguiendo normas locales de confidencialidad de los datos obtenidos y su alcance con el cumplimiento de la ley 25.326.

Participar no conlleva ningún riesgo y los investigadores no obtendremos ningún beneficio económico del mismo. Probablemente se publicarán los resultados de este estudio, pero en ningún caso se utilizarán los nombres de los participantes, sólo las características de los grupos participantes. El único uso de la información será proporcionar datos para el estudio académico.

Su participación es voluntaria. ¿Aceptaría participar?

Si tiene alguna duda sobre este proyecto, Usted puede comunicarse con María de las Nieves Ganiele al 4959-0200, interno 9265 (días hábiles entre 10-14 horas) o vía mail a maria.ganiele@hospitalitaliano.org.ar

El presente trabajo fue presentado para su aprobación por el Comité de Ética de Protocolos de Investigación Universitario (CEPIU), dependiente del Instituto Universitario del Hospital Italiano. Directora: Karin Kopitowski.

- Correo electrónico: karin.kopitowski@hospitalitaliano.org.ar
- Oficina de atención al público: Pte. Perón 4190, Departamento de Investigación, Primer Piso. Potosí 4265, Teléfono de contacto: 4959-0200
- Número de aprobación de protocolo: 0016-22

Estos comités fueron establecidos para proteger los derechos de los sujetos de investigación.

Le agradezco su participación y quedo a su disposición para sugerencias o comentarios.

Bibliografía

- Abdul Rahim, A. F., & Yusoff, M. S. B. (2016). Validity evidence of a multiple mini interview for selection of medical students: Universiti sains Malaysia experience. *Education in Medicine Journal*, 8(2). <https://doi.org/10.5959/eimj.v8i2.437>
- AERA (American Educational Research Association), APA (American Psychological Association) & NCME (National Council on Measurement in Education). (1999). *The Standards for Educational and Psychological Testing*. Washington: AERA (American Educational Research Association).
- Albanese, M. A., Snow, M. H., Skochelak, S. E., Huggett, K. N., & Farrell, P. M. (2003). Assessing personal qualities in medical school admissions. *Academic Medicine: Journal of the Association of American Medical Colleges*, 78(3), 313–321.
- Cameron, Andrea & Mackeigan, Linda. (2012). Development and Pilot Testing of a Multiple Mini-Interview for Admission to a Pharmacy Degree Program. *American journal of pharmaceutical education*. <https://doi.org/10.5688/ajpe76110>.
- Brailovsky, C. (2001). *Educación médica, evaluación de las competencias. Aportes para un cambio curricular en Argentina*. OPS y Facultad de Medicina, UBA. https://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Aportes+para+un+cambio+curricular+en+#d=gs_cit&u=%2Fscholar%3Fq%3Dinfo%3AKSKoiUbY_WsJ%3Ascholar.google.com%2F%26output%3Dcite%26scirp%3D1%26hl%3Den
- Brailovsky, C., & Centeno, A. M. (2012). Algunas Tendencias Actuales en Educación Médica. *Revista de Docencia Universitaria REDU*, 10(Número especial dedicado a la Docencia en Ciencias de la Salud), 23–33.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. In *Qualitative*

Research in Psychology (Vol. 3, Issue 2, pp. 77–101).

<https://doi.org/10.1191/1478088706qp063oa>

Burgos, L. M., de Lima, A. A., Parodi, J., Costabel, J. P., Ganiele, M. N., Durante, E., Arceo, M. D., & Gelpi, R. (2020). Reliability and acceptability of the multiple mini-interview for selection of residents in cardiology. *Journal of Advances in Medical Education & Professionalism*, 8(1), 25.

Carmines, E. G., Zeller, R. A., & Zeller, R. A. (1979). *Reliability and Validity Assessment*. SAGE Publications.

Carroll Turpin, M. A., Steele, K., Matuk-Villazon, O., Rowland, K., Dayton, C. B., & Horn, K. V. (2021). Rapid Transition to a Virtual Multiple Mini-Interview Admissions Process: A New Medical School's Experience During the COVID-19 Pandemic. *Academic Medicine: Journal of the Association of American Medical Colleges*, 96(8), 1152.

Crowl, T. K. (1996). *Fundamentals of Educational Research*. McGraw-Hill Humanities, Social Sciences & World Languages.

Cunnington, J. P., Neville, A. J., & Norman, G. R. (1996). The risks of thoroughness: Reliability and validity of global ratings and checklists in an OSCE. *Advances in Health Sciences Education: Theory and Practice*, 1(3), 227–233.

Diéguez, M. G. (2020). RESIDENCIAS Y PANDEMIA. *REVISTA ARGENTINA DE MEDICINA*, 8(4). <http://revistasam.com.ar/index.php/RAM/article/view/518>

Dore, K. L., Kreuger, S., Ladhani, M., Rolfson, D., Kurtz, D., Kulasegaram, K., Cullimore, A. J., Norman, G. R., Eva, K. W., Bates, S., & Reiter, H. I. (2010). The reliability and acceptability of the Multiple Mini-Interview as a selection instrument for postgraduate admissions. *Academic Medicine: Journal of the Association of American Medical Colleges*, 85(10 Suppl), S60–S63.

- Dowell, J., Lynch, B., Till, H., Kumwenda, B., & Husbands, A. (2012). The multiple mini-interview in the U.K. context: 3 years of experience at Dundee. *Medical Teacher*, 34(4). <https://doi.org/10.3109/0142159X.2012.652706>
- Downing, S. M. (2003). Validity: on meaningful interpretation of assessment data. *Medical Education*, 37(9), 830–837.
- Durante, E., Arceo, M. D., & Ganiele, M. N. (2017). Las múltiples minientrevistas en la selección de residentes de Medicina Familiar. *Investigación En Educación Médica*, 6(22), 137.
- Duré, I. (2015). Residencias médicas en Argentina 2015. *REVISTA ARGENTINA DE MEDICINA*, 3(5), 2–4.
- Edwards, J. C., Johnson, E. K., & Molidor, J. B. (1990). The interview in the admission process. *Academic Medicine: Journal of the Association of American Medical Colleges*, 65(3). <https://doi.org/10.1097/00001888-199003000-00008>
- Eignor, D. R. (2013). The standards for educational and psychological testing. In *APA handbook of testing and assessment in psychology, Vol. 1: Test theory and testing and assessment in industrial and organizational psychology*. (pp. 245–250). <https://doi.org/10.1037/14047-013>
- Eley, D., Young, L., & Przybeck, T. R. (2009). Exploring the temperament and character traits of rural and urban doctors. *The Journal of Rural Health: Official Journal of the American Rural Health Association and the National Rural Health Care Association*, 25(1), 43–49.
- Eva, K. W. (2003). On the generality of specificity. *Medical Education*, 37(7). <https://doi.org/10.1046/j.1365-2923.2003.01563.x>
- Eva, K. W., Macala, C., & Fleming, B. (2019). Twelve tips for constructing a multiple mini-interview. *Medical Teacher*, 41(5).

<https://doi.org/10.1080/0142159X.2018.1429586>

- Eva, K. W., Reiter, H. I., Trinh, K., Wasi, P., Rosenfeld, J., & Norman, G. R. (2009). Predictive validity of the multiple mini-interview for selecting medical trainees. *Medical Education*, 43(8), 767–775.
- Eva, K. W., Rosenfeld, J., Reiter, H. I., & Norman, G. R. (2004). An admissions OSCE: the multiple mini-interview. In *Medical Education* (Vol. 38, Issue 3, pp. 314–326). <https://doi.org/10.1046/j.1365-2923.2004.01776.x>
- Fortin, Y., Kealey, L., Slade, S., & Hanson, M. D. (2016). Investigating Canadian medical school attrition metrics to inform socially accountable admissions planning. *Medical Teacher*, 38(3), 286–290.
- Frank, J. R., Snell, L., & Sherbino, J. (2016). *CanMEDS 2015 Physician Competency Framework*. Ottawa, Canada: Royal College of Physicians and Surgeons of Canada, 2015.
- Gorsuch, R. L. (1997). Exploratory factor analysis: its role in item analysis. *Journal of Personality Assessment*, 68(3), 532–560.
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. SAGE.
- Hair, J., Black, W., Anderson, R., & Babin, B. (2018). *Multivariate Data Analysis, 8th edition*. CENGAGE INDIA.
- Hamdy, H., Prasad, K., Anderson, M. B., Scherpbier, A., Williams, R., Zwierstra, R., & Cuddihy, H. (2006). BEME systematic review: predictive values of measurements obtained in medical schools and future performance in medical practice. *Medical Teacher*, 28(2), 103–116.
- Hammersley, M. (2018). *What's wrong with ethnography?: Methodological explorations*. Routledge.
- Hattie, J. (1985). Methodology Review: Assessing Unidimensionality of Tests and

- Itenls. *Applied Psychological Measurement*, 9(2), 139–164.
- Hecker, K., Donnon, T., Fuentealba, C., Hall, D., Illanes, O., Morck, D. W., & Muelling, C. (2009). Assessment of applicants to the veterinary curriculum using a multiple mini-interview method. *Journal of Veterinary Medical Education*, 36(2). <https://doi.org/10.3138/jvme.36.2.166>
- Hecker, K., & Violato, C. (2011). A generalizability analysis of a veterinary school Multiple Mini Interview: effect of number of interviewers, type of interviewers, and number of stations. *Teaching and Learning in Medicine*, 23(4), 331–336.
- Hofmeister, M., Lockyer, J., & Crutcher, R. (2008). The acceptability of the multiple mini interview for resident selection. *Family Medicine*, 40(10), 734–740.
- Ingreso a residencias*. (2020a). Argentina.gob.ar.
<https://www.argentina.gob.ar/salud/residencias/ingreso-residencias/concurso-unificado>
- Ingreso a residencias*. (2020b, November 3). Argentina.gob.ar.
<https://www.argentina.gob.ar/salud/residencias/ingreso>
- Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, 35(4), 401–415.
- Kane, M. T. (1992). The assessment of professional competence. *Evaluation & the Health Professions*, 15(2), 163–182.
<http://doi.org/10.1177/016327879201500203>
- Klusmann, D., Knorr, M., & Hampe, W. (2022). Exploring the relationships between first impressions and MMI ratings: a pilot study. *Advances in Health Sciences Education: Theory and Practice*. <https://doi.org/10.1007/s10459-022-10151-5>
- Knorr, M., & Hissbach, J. (2014). Multiple mini-interviews: same concept, different approaches. *Medical Education*, 48(12). <https://doi.org/10.1111/medu.12535>
- Kok, K. Y., Chen, L., Idris, F. I., Mumin, N. H., Ghani, H., Zulkipli, I. N., & Lim, M. A.

- (2021). Conducting multiple mini-interviews in the midst of COVID-19 pandemic. *Medical Education Online*, 26(1), 1891610.
- Kornblit, A. L., & Beltramino, F. G. (2004). *Metodologías cualitativas en ciencias sociales: modelos y procedimientos de análisis*. Editorial Biblos.
- Kreiter, C. D., Yin, P., Solow, C., & Brennan, R. L. (2004). Kreiter CD, Yin P, Solow C, Brennan RL. Investigating the reliability of the medical school admissions interview. *Advances in Health Sciences Education: Theory and Practice*, 9(2), 147–159.
- Kulatunga-Moruzi, C., & Norman, G. R. (2002). Validity of admissions measures in predicting performance outcomes: the contribution of cognitive and non-cognitive dimensions. *Teaching and Learning in Medicine*, 14(1), 34–42.
- Kumar, K., Roberts, C., Rothnie, I., du Fresne, C., & Walton, M. (2009). Experiences of the multiple mini-interview: a qualitative analysis. *Medical Education*, 43(4), 360–367.
- 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.
- León, A. B. (2006). La unidimensionalidad de un instrumento de medición: perspectiva factorial. *Revista de Psicología*, 24(1), 53–80.
- Manuel, R. S., Dickens, L., & Young, K. (2019). Qualitative Analysis of Multiple Mini Interview Interviewer Comments. *Medical Science Educator*, 29(4), 941–945.
- Maxwell, J. A. (2012). *Qualitative Research Design: An Interactive Approach*. SAGE Publications.
- Maxwell, J. A. (2022). Interactive Approaches to Qualitative Research Design. In *The SAGE Handbook of Qualitative Research Design* (pp. 41–54).

<https://doi.org/10.4135/9781529770278.n4>

Melro, C. M., Abraham, Z., & Burm, S. (2022). Seven ways to get a grip on preparing for and executing an inclusive virtual multiple mini interview. *Canadian Medical Education Journal*, 13(2), 77–81.

Messick, S. (1989). Meaning and Values in Test Validation: The Science and Ethics of Assessment. In *Educational Researcher* (Vol. 18, Issue 2, pp. 5–11).

<https://doi.org/10.3102/0013189x018002005>

Messick, S. (1995). Standards of Validity and the Validity of Standards in Performance Assessment. In *Educational Measurement: Issues and Practice* (Vol. 14, Issue 4, pp. 5–8). <https://doi.org/10.1111/j.1745-3992.1995.tb00881.x>

Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine: Journal of the Association of American Medical Colleges*, 65(9 Suppl), S63–S67.

Ministerio de Salud de la Nación y Desarrollo Social. Dirección Nacional de Capital Humano. (2019). *La Formación de los Recursos Humanos de Salud. 2016-2019* (García Dieguez Marcelo, L'hospital Cintia, Ladenheim Roberta (Ed.)). Ministerio de Salud de la Nación. Observatorio Federal de Recursos Humanos en Salud.

Nunnally, J., Bernstein, J (1995). *Teoría Psicométrica*. México: McGraw-Hill.

Oliver, T., Hecker, K., Hausdorf, P. A., & Conlon, P. (2014). Validating MMI scores: are we measuring multiple attributes? *Advances in Health Sciences Education: Theory and Practice*, 19(3), 379–392.

Organización Panamericana de la Salud. Área de Sistemas de Salud basados en la Atención Primaria de Salud. (2011). *Residencias médicas en América Latina*. Washington, D.C.: OPS, 2011.

<https://www.paho.org/hq/dmdocuments/2012/HRS-SerieAPSN05-Residencias.p>

df

Patterson, F., & Ferguson, E. (2010). Selection for Medical Education and Training.

In T. Swanwick (Ed.), *Understanding Medical Education* (pp. 352–365).

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). <https://doi.org/10.1111/medu.12817>

Pau, A., Chen, Y. S., Lee, V. K. M., Sow, C. F., & De Alwis, R. (2016). What does the

multiple mini interview have to offer over the panel interview? *Medical Education*

Online, 21. <https://doi.org/10.3402/meo.v21.29874>

Pau, A., Jeevaratnam, K., Chen, Y. S., Fall, A. A., Khoo, C., & Nadarajah, V. D.

(2013). The Multiple Mini-Interview (MMI) for student selection in health

professions training - a systematic review. *Medical Teacher*, 35(12), 1027–1041.

Prideaux, D., Roberts, C., Eva, K., Centeno, A., McCrorie, P., McManus, C.,

Patterson, F., Powis, D., Tekian, A., & Wilkinson, D. (2011). Assessment for

selection for the health care professions and specialty training: consensus

statement and recommendations from the Ottawa 2010 Conference. *Medical*

Teacher, 33(3). <https://doi.org/10.3109/0142159X.2011.551560>

Rees, E. L., Hawarden, A. W., Dent, G., Hays, R., Bates, J., & Hassell, A. B. (2016).

Evidence regarding the utility of multiple mini-interview (MMI) for selection to

undergraduate health programs: A BEME systematic review: BEME Guide No.

37. *Medical Teacher*, 38(5). <https://doi.org/10.3109/0142159X.2016.1158799>

Reiter, H., & Eva, K. (2018). Vive la difference: the freedom and inherent

responsibilities when designing and implementing multiple mini-Interviews.

Academic Medicine: Journal of the Association of American Medical Colleges,

93(7), 969–971.

- Roberts, C., Khanna, P., Rigby, L., Bartle, E., Llewellyn, A., Gustavs, J., Newton, L., Newcombe, J. P., Davies, M., Thistlethwaite, J., & Lynam, J. (2018). Utility of selection methods for specialist medical training: A BEME (best evidence medical education) systematic review: BEME guide no. 45. *Medical Teacher*, 40(1), 3–19.
- Rojas, E. M. (2013). Referentes conceptuales y metodológicos sobre la noción moderna de validez de instrumentos de medición: implicaciones para el caso de personas con necesidades educativas especiales. *Actualidades En Psicología*, 27(114), 113–128.
- Sebok, S. S., Luu, K., & Klinger, D. A. (2014). Psychometric properties of the multiple mini-interview used for medical admissions: findings from generalizability and Rasch analyses. *Advances in Health Sciences Education: Theory and Practice*, 19(1). <https://doi.org/10.1007/s10459-013-9463-7>
- Serralunga, M. G., del Valle, M., Yulita, H., Rodriguez, P., Molina, M. P., Pikas, F., Durante, E., & Garcia Dieguez, M. (2017). Selección de residentes en Argentina: análisis y comparación de los exámenes de admisión de los tres sistemas públicos con mayor oferta de cargos. *Investigación en Educación Médica*, 6(22), e29.
- Singh, N., DeMesa, C., Pritzlaff, S., Jung, M., & Green, C. (2021). Implementation of Virtual Multiple Mini-Interviews for Fellowship Recruitment. *Pain Medicine*, 22(8), 1717–1721.
- Siu, E., & Reiter, H. I. (2009). Overview: what's worked and what hasn't as a guide towards predictive admissions tool development. *Advances in Health Sciences Education: Theory and Practice*, 14(5). <https://doi.org/10.1007/s10459-009-9160-8>

- Sklar, M. C., Eskander, A., Dore, K., & Witterick, I. J. (2015). Comparing the traditional and Multiple Mini Interviews in the selection of post-graduate medical trainees. *Canadian Medical Education Journal*, 6(2), e6–e13.
- St-Onge, C., Côté, D. J., & Brailovsky, C. (2014). Utilisation du Mini Entrevues Multiples en contexte francophone. *Mesure et evaluation en education*, 32(2), 49–69.
- Strauss, A., & Corbin, J. (2002). *Bases de la investigación cualitativa: técnicas y procedimientos para desarrollar la teoría fundamentada*. Universidad de Antioquia.
- Streiner, D. L. (1994). Figuring out factors: the use and misuse of factor analysis. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie*, 39(3), 135–140.
- Streinmer, D. L. (1993). A Checklist for Evaluating the Usefulness of Rating Scales. In *The Canadian Journal of Psychiatry* (Vol. 38, Issue 2, pp. 140–148). <https://doi.org/10.1177/070674379303800214>
- Van Der Vleuten, C. P. (1996). The assessment of professional competence: Developments, research and practical implications. *Advances in Health Sciences Education: Theory and Practice*, 1(1), 41–67.
- Wood, T. J. (2014). Exploring the role of first impressions in rater-based assessments. *Advances in Health Sciences Education: Theory and Practice*, 19(3), 409–427.
- Yusoff, M. S. B. (2019). Multiple Mini Interview as an admission tool in higher education: Insights from a systematic review. *Journal of Taibah University Medical Sciences*, 14(3), 203–240.
- Zaidi, N. L. B., Swoboda, C. M., Kelcey, B. M., & Stephen Manuel, R. (2016). Hidden

item variance in multiple mini-interview scores. *Advances in Health Sciences Education: Theory and Practice*, 22(2), 337–363.

Zúñiga Brenes, M. E., & Rojas, E. M. (2011). Teoría G: un futuro paradigma para el análisis de pruebas psicométricas. In *Actualidades en Psicología* (Vol. 21, Issue 108, p. 117). <https://doi.org/10.15517/ap.v21i108.29>

Listado de las tablas y figuras incluidas en el texto de la tesis por orden de aparición

Tabla 1. Datos demográficos de postulantes

Figura 1. Modelo del esquema de la rotación a través del circuito

Tabla 2. Media y desvío estándar de las puntuaciones por estación

Tabla 3. Varianza total explicada

Tabla 4. Matriz de los componentes

Figura 2. Mapa temático

Tabla 5. Citas de evaluadores organizadas por temas

Tabla 6. Correlación Ítem/ ítem total

Tabla 7. Análisis de la varianza para el modelo PxE

Tabla 8. Tasa de respuesta por especialidad al cuestionario de aceptabilidad

Tabla 9. Respuestas de postulantes al cuestionario de aceptabilidad

Tabla 10. Respuestas de evaluadores al cuestionario de aceptabilidad

Tabla 11. Ventajas y desventajas: selección de comentarios escritos por postulantes

Figura 3: Nube de palabras con los términos más frecuentes referidos por postulantes en sus comentarios escritos

Tabla 12. Ventajas y desventajas: selección de comentarios escritos por evaluadores