Evidence-Based Practices for Interviewing Graduate Medical Education Applicants: A Systematic Review

John C. Lin, ScB
Daniel J. Hu, AB
Ingrid U. Scott, MD, MPH
Paul B. Greenberg, MD, MPH

ABSTRACT

Background Although the selection interview is a standard admission practice for graduate medical education (GME) programs in the United States, there is a dearth of recent reviews on optimizing the trainee interview process, which has low reliability, high cost, and major risk of bias.

Objective To investigate the evidence base for different selection interview practices in GME.

Methods We searched 4 literature databases from inception through September 2022. Two investigators independently conducted title/abstract screening, full-text review, data extraction, and quality assessment. Disagreements were mediated by discussion. We used backward reference searching of included articles to identify additional studies. We included studies of different interview methods and excluded literature reviews, non-GME related publications, and studies comparing different applicant populations. We examined study characteristics, applicant and interviewer preferences, and interview format. We evaluated study quality using the Medical Education Research Study Quality Instrument (MERSQI).

Results Of 2192 studies, 39 (2%) met our inclusion criteria. The evidence base was rated as moderately low quality using MERSQI criteria. Applicants reported preferences for several one-on-one interviews lasting 15 to 20 minutes, interviews by current trainees, and interviews including social events with only trainees. Applicants had mixed perceptions of virtual versus in-person interviews and reported that virtual interviews saved costs. The multiple mini interview (MMI) required more applicant and interviewer time than individual interviews but demonstrated construct and predictive validity and was preferred by applicants and interviewers.

Conclusions Based on moderately low-quality evidence, using the MMI, training interviewers, and providing applicants with basic program information in advance should be considered for GME selection interviews.

Introduction

The selection interview is a standard practice for admission to graduate medical education (GME) programs in the United States.[1] Program directors (PDs) consider the interview important, allowing for the assessment of noncognitive traits and "fit" in applicants.[2] However, there remain significant challenges with the selection interview, including low reliability,[3] high cost,[4] and major risk of bias.[5,6]

Improving the interview process is important to optimize the trainee selection process and match outcomes. Recent reviews studied evidence-based practices for admissions interviews in medical schools[7] and evaluated applicant and program perceptions of virtual interviews in GME.[4] In this systematic review, we investigated the evidence base for different methods of conducting selection interviews in GME, with the goal of identifying best practices.

DOI: http://dx.doi.org/10.4300/JGME-D-23-00115.1

Editor's Note: The online supplementary data contains the literature search strategy used in the study.

Methods

We conducted a systematic review to evaluate current interviewing practices and identify areas for future research.[8] We chose a systematic review because it is a rigorous and comprehensive method to collect, analyze, and synthesize existing evidence on a topic. We developed a literature search strategy (provided as online supplementary data) with a Brown University health sciences librarian. We searched the Embase, ERIC, PubMed, and Web of Sciences databases from inception through September 30, 2022. Two investigators (J.L., D.H.) independently conducted title/abstract screening, full-text review, and data extraction in Covidence in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[9] We evaluated study quality using the Medical Education Research Study

Quality Instrument (MERSQI), which is a validated tool designed to assess the methodological quality of medical education studies across 6 domains: study design, sampling, type of data, validity of evaluation instrument, data analysis, and outcomes.[10] In the literature, MERSQI scores over 12.5 have been considered higher-quality studies,[7] although the MERSQI does not endorse a binary concept of study quality.[10] Disagreements were resolved first by discussion with the 2 investigators, followed by mediation via the senior investigator (P.B.G.) if necessary. Cohen's kappa was calculated to quantify interrater reliability. We used backward reference searching by reviewing the references of included articles to identify additional eligible studies.

We included studies that compared different methods of conducting an admissions interview in GME, including residency and fellowship programs for physician training after medical school. We excluded publications that: (1) were literature reviews, perspectives, or case reports; (2) were not related to GME (eg, in undergraduate medical education); (3) compared interview methods in applicants from different populations (eg, 2 different GME programs) or from different time periods (eg, 1989 vs 2001), as these differences could confound study findings. We included studies conducted at multiple institutions if all applicants were exposed to both the intervention and comparator.

Crossover studies were defined as those in which 2 or more different treatments were applied to subjects at different time periods.[11] Parallel studies were defined as those in which 2 or more groups of subjects received different treatments.[11] Structured interviews were defined as those in which interviewees responded to the same set of mandatory questions or prompts. Semistructured interviews utilized mandatory prompts but allowed interviewers to ask additional questions not provided by programs. Unstructured interviews were defined as those in which the interviewers were not required to ask specific prompts.

The research process included regular team reflections during data collection and analysis. We extracted data on institution, country, publication year, study design, interview methods, study outcomes, and key findings. We then compiled a narrative summary of our results based on the breadth and variety of methods and outcomes included in the review. We identified areas across interviewing that have received significant attention, as well as those that received less. This systematic review was conducted in accordance with the ethical principles of the Declaration of Helsinki.

Results

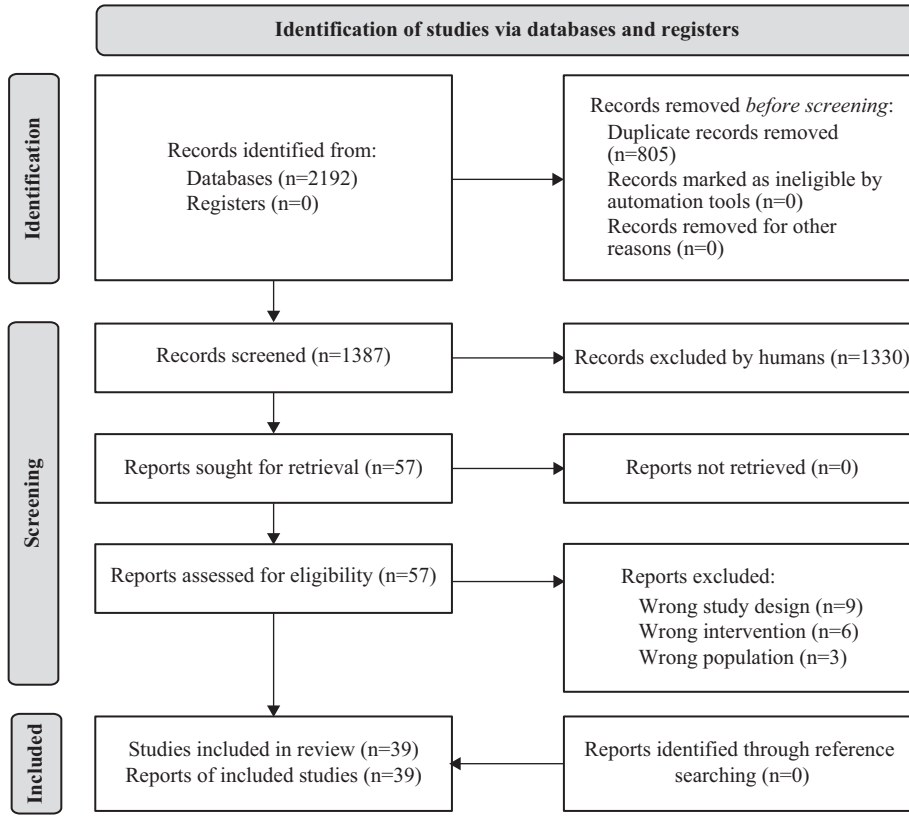
In total, 39 studies met the inclusion criteria after duplicate removal, title/abstract screening, and full-text review (FIGURE).[12-50] Backward reference searching did not identify additional eligible studies. In title/abstract screening, Cohen's kappa was 0.39, and there was 7% disagreement (94 of 1387). In full-text review, Cohen's kappa was 0.37, and there was 22% disagreement (12 of 55). All disagreements were resolved by discussion.

Study Characteristics

The 39 included studies are described in TABLE 1. The studies were published from 1985 to 2022, with 37 (95%) studies published after 2000, and 32 (82%) studies published from 2010 to 2022. All but one were conducted in World Bank-classified high-income countries: 24 (62%) in the United States, 8 (21%) in Canada, 3 (8%) in Japan, and 1 (3% each) in Argentina, Denmark, Oman, and the Netherlands. Twenty-six studies (67%) were conducted at a single institution, 3 (8%) at multiple institutions, 2 (5%) did not report their institution, and 8 (21%) were general surveys sent to many institutions. Thirty-five studies (90%) evaluated residency programs, spanning over 19 distinct specialties. Four (10%) evaluated fellowship programs, covering 3 specialties. Additionally, 9 (23%) were crossover studies, 6 (15%) were parallel studies, 7 (18%) were post-interview surveys, 3 (8%) were post-application surveys, 3 (8%) were crossover studies with post-interview surveys, 2 (5%) were randomized controlled trials, and 1 (3%) was a pre-interview survey. The included studies were rated as moderately low quality by MERSQI (TABLE 2), averaging a mean (standard deviation [SD]) of 10.9 (1.5) out of a maximum of 18.

Thirty-four of the 39 studies (87%) reported their sample size. Of these, the mean sample was 187.7 (SD=338.1). Response rates of the 14 surveys (36%) ranged from 21% to 100%. Ten studies (26%) compared in-person to virtual interviews; 8 (21%) compared unstructured interviews to the multiple mini interviews (MMIs), which involve several short, separate interviews by different interviewers for each applicant[20]; 6 (15%) compared more than 2 interview methods, 4 (10%) compared different MMI formats, 4 (10%) compared structured to unstructured interviews, 3 (8%) compared closed-file to open-file interviews, and 2 (5%) compared semi-structured to unstructured interviews. The remaining studies compared faculty to resident interviewers,[16] 2 to 3 interviewers,[47] informal to formal dress codes,[31] use of a pre-interview informational video to no pre-interview video,[27] and a single-site interview fair to individual program interviews.[28]

Twelve outcomes were identified from included studies: applicant costs,[12,28,31,40,41] applicant perceptions,[12,13,17-22,28,31,33-35,37,40,41,45,49] applicant time,[27,38]



FIGURE

PRISMA Flow Diagram

Note: We incorporated studies examining various interview methodologies within graduate medical education (GME), encompassing both residency and fellowship programs post-medical school. Exclusions were: (1) literature reviews, perspectives, and case reports; (2) studies outside the GME purview such as undergraduate medical education; (3) those contrasting disparate applicant cohorts; and (4) studies that did not compare interventions. Notably, during our full-text review, a significant number of excluded studies were either single-arm investigations or pertained to undergraduate medical education.

concurrent validity,[38] construct validity,[36,43,44] institutional costs,[15,28,39] interview bias,[26,30,42] interview reliability,[24,25,29,32,47,49,50] interviewer time,[15,23,27,38,39] interviewer perceptions,[19,28,41,49,50] PD perceptions,[15,34,37,39,46] and predictive validity.[14,16,48]

Study Outcomes

Applicant Preferences: In surveys sent to applicants by individual programs in gastroenterology, orthopedics, and urology, 67% of applicants preferred one-on-one (as opposed to 1:2, 1:3, 1:4) interviews,[34] 55 to 68% preferred 15-to-20 minute interviews,[18,34] 83% preferred 5- to 7-minute office setting faculty interviews,[33] and 50 to 95% preferred interview days that lasted one-half to three-fourths of a workday.[18-33] Ninety-six percent of orthopedic respondents wanted to be interviewed by current residents, and 94% wanted to have a social event in the interview process;[18] 88% of urology applicants wanted to interview with at least half of a program's faculty.[33] Of orthopedic applicants, 36% wanted skills tasks and 23% wanted knowledge tests during their interview.[18] Ninety-seven percent of urology applicants believed that interview offers should be released on the same day.[20]

In one study, 85% of applicants surveyed preferred an explicitly informal dress code over an unspoken formal dress code; 22% of respondents reported that this led them to rank the program higher and 28% reported it reduced their costs.[31] Showing applicants a video about program information before the interview reduced time spent on program information during the interview.[27] A post-interview survey conducted by Canadian urology programs showed that applicants and interviewers preferred a single-site, single-day interview fair with multiple programs present over individual interviews in multiple geographically distinct sites with each program/applicant; applicant costs were reduced from an estimated $2,065 to $367 for the fair, although institutional costs were slightly higher ($1,931 rather than an estimated $1,825).[28]

TABLE 1
Design and Details of Included Studies

Study (Year)[a]	Institution	Specialty	Program Type	Country	Design	Comparator
Greene (1985)[27]	University of Massachusetts	Internal medicine	Residency	United States	RCT	Pre-interview videotape vs without
Altmaier (1992)[14]	University of Iowa	Radiology	Residency	United States	Crossover[b]	Behavioral vs traditional
Gilbart (2001)[25]	University of Toronto	Orthopedic surgery	Residency	Canada	Parallel[c]	Unstructured vs semistructured[d]
Grober (2003)[28]	Dalhousie University, McGill University, Queen's University, University of Alberta, University of British Columbia, University of Manitoba, University of Ottawa, University of Toronto, University of Western Ontario	Urology	Residency	Canada	Post-interview survey	Single-site "interview fair" vs individual program interviews
Swanson (2005)[44]	Children's Hospital of Philadelphia	Pediatrics	Residency	United States	Crossover	Closed vs open file (USMLE/grades)
Hamel (2007)[29]	Université de Montréal	Ophthalmology	Residency	Canada	Crossover	Different interview topics
Hauge (2007)[30]	Rush University Medical Center	General surgery	Residency	United States	Crossover	Closed vs open file
Dore (2010)[22]	McMaster University, University of Alberta	OB/GYN, pediatrics, internal medicine	Residency	Canada	Post-interview survey	MMI vs traditional
Finlayson (2011)[23]	University of British Columbia	PM&R	Residency	Canada	Parallel	MMI vs traditional
Shah (2012)[41]	University of New Mexico	Urology	Residency	United States	Crossover, post-interview survey	In-person vs virtual
Isaksen (2013)[32]	University of Southern Denmark	Family medicine	Residency	Denmark	Crossover	Unstructured vs semistructured interviews
Lubarsky (2013)[36]	McGill University	Neurology	Residency	Canada	Crossover	Different MMI stations
Vermeulen (2013)[47]	University Medical Centre Utrecht	Family medicine	Residency	The Netherlands	Parallel	2 vs 3 semistructured interviewers
Campagna-Vaillancourt (2014)[19]	McGill University	ENT	Residency	Canada	Post-interview survey	MMI vs traditional
Jacobs (2014)[33]	Medical College of Wisconsin	Urology	Residency	United States	Pre-interview survey	Different interview methods
Hern (2015)[31]	Alameda Health System	Emergency medicine	Residency	United States	Post-interview survey	Informal vs formal dress code
Sklar (2015)[43]	University of Toronto	ENT	Residency	Canada	Crossover	MMI vs unstructured
Yoshimura (2015)[50]	Tokyo Bay Urayasu Ichikawa Medical Center	Internal medicine, surgery, emergency medicine	Residency	Japan	Crossover, post-interview survey	MMI vs traditional; behavioral vs situational MMI questions
Camp (2016)[18]	Mayo Clinic, Orlando Health, University of Iowa, University of Southern California	Orthopedic surgery	Residency	United States	Post-application survey	Different interview methods
Ogunyemi (2016)[38]	Cedars-Sinai Medical Center	OB/GYN	Residency	United States	Parallel	MMI vs traditional, different MMI methods
Vadi (2016)[45]	Loma Linda University	Anesthesiology	Residency	United States	Post-interview survey	In-person vs virtual
Shipper (2017)[42]	Stanford University	General surgery	Residency	United States	Crossover	Closed vs open file
VanOrder (2017)[46]	Michigan State University	All	Residency	United States	General survey	Different interview methods

TABLE 1
Design and Details of Included Studies (continued)

Study (Year)[a]	Institution	Specialty	Program Type	Country	Design	Comparator
Yamada (2017)[49]	Tokyo Bay Urayasu Ichikawa Medical Center	Internal medicine, surgery, emergency medicine	Residency	Japan	Crossover, post-interview survey	Behavioral vs situational MMI questions
Gardner (2018)[24]	NR	Minimally invasive and bariatric surgery	Fellowship	United States	Parallel	Structured vs unstructured
Al Abri (2019)[13]	Sultan Qaboos University	ENT	Residency	Oman	Post-interview survey	MMI vs traditional
Beran (2019)[16]	Medical College of Wisconsin	OB/GYN	Residency	United States	Crossover	Unstructured faculty vs structured faculty vs unstructured residents
Burgos (2020)[17]	Instituto Cardiovascular de Buenos	Cardiology	Residency	Argentina	Post-interview survey	MMI vs other types
Gordon (2020)[26]	University of Pennsylvania	Anesthesiology	Residency	United States	RCT	Behavioral vs unstructured
Kenigsberg (2020)[35]	UT Southwestern	Urology	Residency	United States	Post-application survey	In-person vs virtual
Seifi (2020)[40]	UT Health San Antonio residency programs and Texas medical schools	All	Residency	United States	General survey	In-person vs virtual
Carpinito (2021)[20]	UT Southwestern	Urology	Residency	United States	Post-application survey	Different interview methods
D'Angelo (2021)[21]	Mayo Clinic	Colorectal surgery	Fellowship	United States	General survey	In-person vs virtual
Kamboj (2021)[34]	Mayo Clinic	Gastroenterology	Fellowship	United States	General survey	In-person vs virtual
Rajesh (2021)[39]	UT Health San Antonio	General surgery	Residency	United States	General survey	In-person vs virtual
Yamada (2021)[48]	NR	NR	Residency	Japan	Parallel	Behavioral vs situational MMI questions
Abelson (2022)[12]	Ohio State University	Colorectal surgery	Fellowship	United States	General survey	In-person vs virtual
Asaad (2022)[15]	Baylor College of Medicine	General, neurologic, OB/GYN, ophthalmic, orthopedic, ENT, plastic, thoracic, and vascular surgery	Residency	United States	General survey	In-person vs virtual
Moran (2022)[37]	33 programs	Radiology	Residency	United States	General survey	In-person vs virtual

Abbreviations: RCT, randomized controlled trial; USMLE, United States Medical Licensing Examination; OB/GYN, obstetrics and gynecology; MMI, multiple mini interview; PM&R, physical medicine and rehabilitation; ENT, ears, nose, and throat; NR, not reported;
[a] Studies are organized by year published. Within each year, articles are organized alphabetically by author name.
[b] Crossover studies were defined as those in which 2 or more different treatments were applied to subjects at different time periods.
[c] Parallel studies were defined as those in which 2 or more groups of subjects received different treatments.
[d] Unstructured interviews were defined as those in which the interviewers were not required to ask specific prompts. Structured interviews were defined as those in which interviewers asked only the same set of mandatory questions or prompts, whereas semistructured interviews utilized mandatory prompts but allowed interviewers to ask additional questions not provided by programs.

Interviewer Practices: Interviewers rated applicants higher based on their grades, United States Medical Licensing Examination (USMLE) scores, Alpha Omega Alpha status, school ranking, and application score, when available.[42,44] Interviewers initially rated applicants who had rotated in their department higher, although this effect disappeared after they reviewed applicant files in one study.[30] Interview ratings for the communication, collaboration, management, medical expertise, and scholarship domains had acceptable reliability (intraclass correlations: 0.900, 0.739, 0.585, and 0.585, respectively) in a Canadian ophthalmology program, but ratings for the professionalism and health advocacy domains did not.[29]

TABLE 2
Findings and Outcomes of Included Studies

Study (Year)[a]	Sample Size	RR	MERSQI (Out of 18)	Variables of Interest	Findings
Greene (1985)[27]	77	NR	11.5	Applicant time, interviewer time	Showing applicants a videotape on basic information reduced time spent on basic info from 7.4 to 1.3 minutes.
Altmaier (1992)[14]	NR	NR	8.5	Predictive validity	Unstructured faculty and resident interviews and structured behavioral interviews did not predict a resident's clinical ratings (by their PD). "Interpersonal skills" rating of behavioral interviews was correlated.
Gilbart (2001)[25]	66	NR	11.5	Interview reliability	There were no differences in interview reliability between unstructured and semistructured interviews across programs.
Grober (2003)[28]	36	97%	11	Applicant perceptions, applicant costs, interviewer perceptions, institutional costs	Nearly all (96%) of applicants preferred the interview fair format. Applicant costs for the fair were $367 (compared to an estimated $2,065 for individual interviewing). All interviewers preferred the interview fair format. Institutional costs were $1,931 (compared to an estimated $1,825).
Swanson (2005)[44]	935	NR	12	Construct validity	When interviewers could see USMLE scores and grades, interviews were correlated with USMLE scores.
Hamel (2007)[29]	25	NR	11.5	Interview reliability	Interrater reliability was acceptable for interviews about communication, collaboration, management, medical expertise, and scholarship, but not about professionalism or health advocacy.
Hauge (2007)[30]	70	NR	12	Interview ratings	Interviewers rated applicants who had rotated in their department higher, but this effect disappeared after reviewing applicant files.
Dore (2010)[22]	NR	NR	12.5	Applicant perceptions	74% of respondents reported that MMI was better than traditional interviews.
Finlayson (2011)[23]	NR	NR	11.5	Interviewer time	MMI required 4 hours for 16 applicants, whereas prior panel interviews required 16 hours.
Shah (2012)[41]	37	95%	12	Applicant perceptions, applicant costs, interviewer perceptions	Applicants and interviewers preferred continuing virtual interviews in addition to in-person interviews. Applicants spent less money ($171 vs $364) and time away from school with virtual interviews. Applicants perceived virtual interviews as less effective.
Isaksen (2013)[32]	29	NR	11	Interview reliability	Semistructured interviews were reliable in 1 out of 2 rounds.

TABLE 2
Findings and Outcomes of Included Studies (continued)

Study (Year)[a]	Sample Size	RR	MERSQI (Out of 18)	Variables of Interest	Findings
Lubarsky (2013)[36]	29	94%	9.5	Construct validity	Only 3 of 28 (11%) correlations between global rating scores were significant, indicating that MMI stations measured different constructs.
Vermeulen (2013)[47]	365	NR	11.5	Interview reliability	Reliability decreased from an average of 0.82 to 0.72 when 2, rather than 3, interviewer ratings were used.
Campagna-Vaillancourt (2014)[19]	63	100%	11.5	Applicant perceptions, interviewer perceptions	73% of applicants and 78% of interviewers agreed that they preferred the MMI over traditional interviews.
Jacobs (2014)[33]	221	100%	9	Applicant perceptions	80+% of applicants wanted: 5-7 faculty interviews, interviewing with 50-75% of department faculty, in office settings, and lasting 1/2-3/4 of a workday.
Hern (2015)[31]	236	76.6%	9	Applicant perceptions, applicant costs	85.1% of applicants preferred an explicitly informal dress code; 27.7% said that this reduced their costs; 21.7% said they ranked the program higher.
Sklar (2015)[43]	27	NR	12	Construct validity	Although the MMI and unstructured interview were moderately correlated, there was very poor agreement on rank.
Yoshimura (2015)[50]	26	100%	11.5	Applicant perceptions, interviewer perceptions, interview reliability	All applicants and interviewers agreed that the MMI was fairer than the traditional interview; most applicants believed that situational questions assessed candidates the best, whereas most interviewers felt that behavioral questions assessed candidates the best. Reliability was acceptable for both behavioral and situational MMI stations.
Camp (2016)[18]	408	37%	9	Applicant perceptions	94% of applicants wanted a social event in the interview process; 54% wanted a social event with only residents. Only 36% wanted manual dexterity or skills tasks, and only 23% wanted orthopedic knowledge tests. 87.9% wanted interviews to last a half day. 55% wanted interviews to last 15 minutes; 33% preferred 20-minute interviews. On average, applicants wanted 5 interviews total, with 1.7 interviewers per interview. 96% wanted to be interviewed by current residents. Applicants said that the interview was important for their rank lists.

TABLE 2
Findings and Outcomes of Included Studies (continued)

Study (Year)[a]	Sample Size	RR	MERSQI (Out of 18)	Variables of Interest	Findings
Ogunyemi (2016)[38]	199	NR	11.5	Concurrent validity, construct validity, applicant time, interviewer time	Behavioral interviews were correlated with USMLE Step 2 scores; role play was correlated with USMLE Step 1 scores; social media ethics and surgical technical simulation were correlated with male sex; MMI stations were correlated with each other. The MMI increased interview day length by 15% and doubled face-to-face interview time for applicants. Faculty/residents spent 23% more time in the MMI.
Vadi (2016)[45]	111	87%	12.5	Applicant perceptions	Applicants who chose virtual interviews did so due to scheduling conflicts, distance, and costs. Applicants who chose in-person interviews did so due to interest in resident interactions, geographic proximity, and interest in visiting campus.
Shipper (2017)[42]	NR	NR	10.5	Interview bias	Open file interviewers rated applicants higher based on their Alpha Omega Alpha status, school ranking, and application score.
VanOrder (2017)[46]	83	44.6%	12	PD perceptions	The type of interview method was not correlated with subsequent satisfaction with selected residents.
Yamada (2017)[49]	40	NR	11.5	Applicant perceptions, interviewer perceptions, interview reliability	Applicants did not significantly prefer either format; interviewers preferred behavioral over situational questions. Behavioral questions had a lower reliability than situational questions.
Gardner (2018)[24]	5	NR	12	Interview reliability	After training, faculty improved their interrater agreement on applicant competency, with 80% of ratings being within 2 points of each other.
Al Abri (2019)[13]	NR	NR	12.5	Applicant perceptions	All (100%) applicants stated that they preferred MMI over traditional interviews and that MMI was more effective than traditional interviews.
Beran (2019)[16]	44	NR	12	Predictive validity	Behavioral-based interviews were not associated with future faculty and nursing evaluations; unstructured faculty interviews were associated with faculty evaluations; resident team interviews were associated with faculty and nursing evaluations.
Burgos (2020)[17]	22	100%	12	Applicant perceptions	91% of applicants preferred MMI over other types of interviews.
Gordon (2020)[26]	187	NR	14.5	Interview bias	Structured behavioral and unstructured interviews were rated similarly.

TABLE 2
Findings and Outcomes of Included Studies (continued)

Study (Year)[a]	Sample Size	RR	MERSQI (Out of 18)	Variables of Interest	Findings
Kenigsberg (2020)[35]	156	24%	8	Applicant perceptions	A plurality (41%) of applicants preferred in-person interviews. Most believed city visits (87%) and resident interactions (81%) could not be replicated virtually, but faculty interviews (81%) could; 64% rated resident interactions as the most important interview day component.
Seifi (2020)[40]	1824	24%	10	Applicant perceptions, applicant costs	Applicants agreed with more statements on the efficacy of in-person interviews; applicants were uneasy about the travel expenses of in-person interviewing.
Carpinito (2021)[20]	144	36%	8	Applicant perceptions	A plurality (39%) of applicants preferred virtual interviews. 97% thought all interview offers should be released on the same day. Faculty and resident interviews had the greatest impact on rank lists. 65% spent <$2,000 for their application cycle.
D'Angelo (2021)[21]	77	54%	9.5	Applicant perceptions	73% of applicants recommended virtual interviews even if COVID-19 was not a factor.
Kamboj (2021)[34]	162	26%	9	Applicant perceptions, PD perceptions	Applicants preferred one-on-one interviews (67%) and 20-minute interviews (68%); programs preferred one-on-one structured interviews (62%) and for 20 minutes (44%). Applicants and programs had mixed perceptions of virtual and in-person interviews, but had the least support for virtual interviews.
Rajesh (2021)[39]	60	21%	9	PD perceptions, interviewer time, institutional costs	93% felt that virtual interviews were less expensive than in-person interviews; only 35% felt they were less time-consuming. 50% did not believe that virtual interviews were overall better. 65% said they will offer virtual and in-person interviews in the future. Average cost savings were $6,462.
Yamada (2021)[48]	25	NR	9.5	Predictive validity	Behavioral and situational MMI questions were both correlated with faculty evaluations after 1 year. Only behavioral questions were correlated with faculty evaluations after 2 years.
Abelson (2022)[12]	119	48%	11	Applicant perceptions, applicant costs	73% of virtual applicants preferred virtual interviews even if COVID-19 was not a factor; 53% of in-person applicants preferred in-person interviews. 53% of in-person applicants spent >$5,000 on interviews.

TABLE 2
Findings and Outcomes of Included Studies (continued)

Study (Year)[a]	Sample Size	RR	MERSQI (Out of 18)	Variables of Interest	Findings
Asaad (2022)[15]	365	25.9%	10	PD perceptions, interviewer time, institutional costs	57% of PDs disagreed that virtual interviews were better than in-person interviews; a plurality (43%) reported that virtual was not less time-consuming than in-person; 90% found virtual interviews to be less expensive than in-person interviews.
Moran (2022)[37]	109	49.5%	10.5	Applicant perceptions, PD perceptions	80% of applicants and 60% of PDs believed that the benefits of virtual interviews outweighed the drawbacks.

Abbreviations: RR, response rate; MERSQI, Medical Education Research Study Quality Instrument; NR, not reported; USMLE, United States Medical Licensing Examination; MMI, multiple mini interview; PD, program director.
[a] Studies are organized by year published. Within each year, articles are organized alphabetically by author name.

Interview Format: In-Person vs Virtual: Ten of the 39 studies (26%) were conducted during or after the onset of the COVID-19 pandemic. Of these, 8 compared in-person versus virtual formats.

Applicants and PDs had mixed perceptions of virtual interviews. In 6 surveys of applicants, 1 reported that 80% of radiology applicants during the COVID-19 virtual interview season agreed that the benefits of the virtual interview season outweighed the drawbacks,[37] 1 reported that 73% of colorectal surgery applicants recommended virtual interviews,[21] 3 surveys reported that gastroenterology and urology applicants had mixed preferences,[20,34,35] and 1 reported that virtual colorectal surgery applicants preferred virtual interviews and in-person applicants preferred in-person interviews.[12] When allowed to choose an interview format, anesthesiology applicants who selected virtual interviews did so due to scheduling conflicts, distance, and costs, whereas applicants who selected in-person interviews did so for resident interactions, geographic proximity, and visiting campus.[45] Of 4 general PD surveys, 2 reported that surgical PDs did not believe that virtual interviews were better than in-person interviews, 1 reported that gastroenterology PDs had mixed perceptions, and 1 reported that radiology PDs preferred virtual interviews. In a post-interview survey, virtual urology applicants and interviewers preferred continuing virtual interviews in addition to in-person interviews.[41]

Applicants and PDs found virtual interviews more affordable than in-person interviews but disagreed on whether they saved time. Three general surveys found that applicants across all specialties on average spent less money for virtual interviews.[12,40,41] On average, general surgery programs saved $6,462 after one year of virtual interviews. A general survey found that for urology applicants, 10% missed days at school with virtual interviews versus 30% with in-person interviews.[41] Surgical PDs did not believe that virtual interviews were less time-consuming for programs.[15,39]

Interview Format—The MMI: Applicants and interviewers reported to programs that they preferred the MMI over unstructured and other types of interviews. Most or all applicants preferred MMIs in 5 post-interview surveys in 7 specialties (100%,[13] 91%,[17] 78%,[19] 74%,[22] 100%[50]), and most interviewers preferred the MMI in 1 post-interview survey.[19] In single studies, the MMI required less faculty time than panel interviews,[23] required more faculty and resident time than individual interviews,[38] extended interview day length by 15%,[38] and doubled face-to-face interview time for applicants.[38] Another study found that the MMI and unstructured interview were moderately correlated but disagreed on rank order lists, indicating that the MMI and unstructured interview measured different constructs.[43]

In MMI stations, behavioral and situational questions demonstrated reliability per generalizability theory, though behavioral questions were preferred by interviewers, and they predicted 2-year faculty evaluations of residents.[49,50] Applicants did not express a preference, but interviewers preferred behavioral over situational questions.[49,50] Behavioral questions had lower interview reliability than situational questions,[49] although both types had acceptable reliability (G=0.87 and 0.96, respectively).[50] One study in neurology found that MMI station ratings were generally not significantly correlated,[36] while another study in obstetrics and gynecology (OB/GYN) found that several MMI station ratings were correlated with each other.[38]

These results suggest that MMI stations overlap in their construct assessment. Behavioral questions were correlated with USMLE Step 2 scores; role play ratings were correlated with USMLE Step 1 scores; and social media ethics and surgical technical simulation ratings were correlated with male sex.[38] Behavioral and situational questions were both correlated with 1-year faculty evaluations of residents in 3 specialties, but only behavioral questions were correlated after 2 years.[48]

Interview Format—Structured (Non-MMI) vs Unstructured: Unstructured, semistructured, and non-MMI structured interviews mostly had similar interrater reliability. A study in orthopedics reported that there were no differences in interrater reliability between semistructured and unstructured interviews;[25] another study in family medicine found that semistructured but not unstructured interviews were reliable in 1 out of 2 rounds.[32] Interview reliability decreased slightly when 2, rather than 3, semistructured interview ratings were considered.[47] However, interviewer training improved interrater reliability.[24] Structured and unstructured interviews in anesthesiology were rated similarly.[26]

The evidence was mixed regarding predictive validity of unstructured and structured interviews. Neither unstructured nor structured behavioral interviews predicted an applicant's future clinical rating by their PD in one radiology residency program.[14] Similarly, another study reported that different interview methods across residency programs were not correlated with subsequent PD satisfaction with their matched residents.[46] In contrast, unstructured faculty and resident team interview scores were associated with future faculty and nursing evaluations in one OB/GYN residency program.[16]

Discussion

In this review of 39 studies of overall moderately low quality, we investigated different selection interview methods in GME to identify evidence-based practices. Applicants had mixed perceptions of virtual versus in-person interviews and generally preferred multiple one-on-one interviews and including interviews with current residents. Both applicants and PDs reported that virtual interviews saved costs. The MMI was preferred by applicants and interviewers over unstructured interviews and demonstrated construct validity, but it required more applicant and interviewer time than individual interviews. Aside from the MMI, adding structure to interviews did not improve reliability or predictive validity in this review. Masking interviewers to applicants' academic records reduced the influence of academic performance on interview scores.

Implications for Interviewing in GME

As advised by the Association of American Medical Colleges, GME programs may consider conducting virtual interviews,[51] which save time and money for applicants and programs and may help make the application process more equitable.[51] Additionally, though there is a paucity of research on the reliability and validity of virtual MMIs in GME,[4] virtual and in-person MMI scores were similar for medical and dental school admissions.[52] Hence, a virtual MMI may be an option for programs seeking to utilize virtual interviews.

Within the 39 studies, conducted mostly in North American, university-based residency programs, the MMI had acceptable reliability, construct validity, predictive validity for 1- and 2-year faculty evaluations of trainees, and was rated favorably by applicants and interviewers in post-interview surveys. These findings align with a systematic review of admissions interviews in undergraduate medical education that determined that the MMI was reliable, unbiased, and predicted first-year performance on written examinations and Objective Structured Clinical Examinations.[8] In one study, behavioral and situational MMI questions were both reliable and predictive of 1-year faculty evaluations of residents, and only behavioral MMI questions were predictive of 2-year faculty evaluations of residents.[48] This information about the reliability and validity of behavioral MMI questions may be helpful when designing MMIs.[48] Another study showed that the MMI and unstructured interviews produced different rank order lists.[43] In comparison to the MMI, there was limited reliability and validity evidence for unstructured and structured interviews. However, potential drawbacks of the MMI include potential interviewer biases towards applicants based on age, rural background, and cultural and language barriers.[53-56] Additionally, these findings may not be applicable to less-studied settings like community-based programs and populations like international medical graduates.

Although the evidence base was weak and meta-analysis was not performed, this review uncovered several general considerations for the conduct of GME interviews. Training interviewers may help ensure impartial, reliable interviewing. Resident team interviews can be predictive of faculty and nursing evaluations; indeed, applicants wanted to be interviewed by current residents. Interviewers provided with academic information before interviews, including USMLE scores and grades, may prejudge applicants accordingly. Additionally, providing basic program information before interviews may allow applicants and interviewers to discuss other information. When deciding upon an

interview format, programs should also identify clear goals for the interview. For example, programs may seek to evaluate applicants' interpersonal and problem-solving skills or assess their awareness and sensitivity toward different cultures and social backgrounds. Programs can also structure their interview day agenda to align with applicant preferences. Based on 3 surveys, these include holding 5 to 7 one-on-one interviews lasting 15 to 20 minutes, separate social events with only residents and with faculty, and an informal dress code. Applicants also preferred not to be tested on their skills or knowledge. However, the costs of implementation may render some of these ideas impractical for many programs.

This systematic review evaluated multiple interview methods in resident and fellow selection, building upon the work of prior reviews[3,5] with a standardized literature search, adherence to a predefined protocol, use of PRISMA reporting guidelines, and study quality assessment with MERSQI. A 2015 literature review found that masking interviewers could reduce interviewer bias and had mixed findings regarding the predictive value of interviews for performance in residency.[2] A 2022 systematic review on virtual interviewing identified applicant and interviewer satisfaction with virtual interviews as well as cost and time savings but did not investigate the predictive validity or reliability of a virtual MMI.[4]

Limitations of the Evidence Base

Current studies on admissions interviewing in GME have several key limitations. There was a paucity of high-quality studies; only 2 randomized controlled trials were eligible for inclusion in this review, and 35 of the 39 included studies (90%) were rated as low quality by MERSQI. Of eligible studies, few were available for many outcomes of interest, and all had been published in or after 1985. Several studies did not report important information, such as institution and specialty. None of the studies compared patient outcomes for admitted residents based on their interview performance; instead, included studies used proxies such as faculty and nursing evaluations. Many studies did not describe their interview procedures, including how the interviewers were selected and trained. There was widespread heterogeneity in interview procedures, survey methodology, and outcome measurement that precluded meta-analysis. Some findings may not be transferable among specialties or institutions, as smaller departments may have fewer resources and faculty for interviews. Additionally, most relevant studies were conducted at academic institutions in our literature search. Hence, our findings may be less generalizable to GME programs—such as community-based programs[57]—underrepresented in published research.[57] More research is needed to identify effective interview methods for these programs.

Moreover, the surveys used to evaluate interviewer and interviewee preferences were subject to several limitations. The program-specific surveys used in 11 (28%) of the included studies were vulnerable to response bias.[58,59] Methods to minimize response bias include highlighting the scientific nature of the survey and its importance for future applicants, avoiding questions that may identify specific respondents, sending the survey after the interview process is completed and stating that responses will only be reviewed post-match.[58-61] However, a general survey of colorectal surgery applicants found that applicants interviewed virtually preferred virtual interviews and that in-person interviewees preferred in-person interviews,[12] which suggests that interview satisfaction surveys may be more broadly subject to response bias. While 29 (74%) of the included studies were conducted before the COVID-19 pandemic, 8 of the 10 studies comparing virtual versus in-person interviews were conducted during or shortly after the COVID-19 public health emergency ended. Applicant and interviewer preferences may shift over time as programs refine the virtual format and explore its implications for the trainee selection process.

Limitations of the Review

There are several limitations to this review. We may have omitted relevant articles. This could arise from errors in our search strategy, although we consulted a health sciences librarian to optimize our search, reviewed the reference lists of included articles, and searched 4 major health and education databases. This could also arise from investigator bias and errors, although 2 investigators independently conducted each step of study identification and data extraction with third party mediation. In addition, we limited the review to comparative studies, which may have omitted related studies but ensured that more rigorous evidence was included.[10] Quantitative meta-analysis was not possible due to study heterogeneity. Finally, we did not study the weighting of the interview in developing rank-order lists in GME, which also impacts the trainee selection process.

Conclusions

Based on moderately low-quality evidence, this systematic review found several practices that should be considered in the GME selection interview, including

using the MMI, conducting interviewer training, and providing applicants with information about a training program in advance.

References

1. National Resident Matching Program. Results of the 2022 NRMP Program Director Survey. Accessed November 16, 2022. https://www.nrmp.org/wp-content/uploads/2022/09/PD-Survey-Report-2022_FINALrev.pdf
2. Stephenson-Famy A, Houmard BS, Oberoi S, Manyak A, Chiang S, Kim S. Use of the interview in resident candidate selection: a review of the literature. *J Grad Med Educ*. 2015;7(4):539-548. doi:10.4300/jgme-d-14-00236.1
3. Kreiter CD, Yin P, Solow C, Brennan RL. Investigating the reliability of the medical school admissions interview. *Adv Health Sci Educ Theory Pract*. 2004;9(2):147-159. doi:10.1023/B:AHSE.0000027464.22411.0f
4. Daniel M, Gottlieb M, Wooten D, et al. Virtual interviewing for graduate medical education recruitment and selection: a BEME systematic review: BEME guide no. 80. *Med Teach*. 2022;44(12):1313-1331. doi:10.1080/0142159X.2022.2130038
5. Leduc J-M, Rioux R, Gagnon R, Bourdy C, Dennis A. Impact of sociodemographic characteristics of applicants in multiple mini-interviews. *Med Teach*. 2017;39(3):285-294. doi:10.1080/0142159X.2017.1270431
6. Chatterjee A, Greif C, Witzburg R, Henault L, Goodell K, Paasche-Orlow MK. US medical school applicant experiences of bias on the interview trail. *J Health Care Poor Underserved*. 2020;31(1):185-200. doi:10.1353/hpu.2020.0017
7. Lin JC, Lokhande A, Margo CE, Greenberg PB. Best practices for interviewing applicants for medical school admissions: a systematic review. *Perspect Med Educ*. 2022;11(5):239-246. doi:10.1007/s40037-022-00726-8
8. McGaghie WC. Varieties of integrative scholarship: why rules of evidence, criteria, and standards matter. *Acad Med*. 2015;90(3):294-302. doi:10.1097/ACM.0000000000000585
9. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. doi:10.1136/bmj.n71
10. Reed DA, Cook DA, Beckman TJ, Levine RB, Kern DE, Wright SM. Association between funding and quality of published medical education research. *JAMA*. 2007;298(9):1002-1009. doi:10.1001/jama.298.9.1002
11. Louis TA, Lavori PW, Bailar JC III, Polansky M. Crossover and self-controlled designs in clinical research. In: Bailar JC III, Mosteller F, eds. *Medical Uses of Statistics*. 2nd ed. CRC Press; 1992:83-104.
12. Abelson JS, Al Jabri A, Kelley SR, D'Angelo JD, D'Angelo AD, Husain S. Applicants' perspectives on virtual vs in-person colorectal residency interviews. *Dis Colon Rectum*. 2022;65(10):1162-1166. doi:10.1097/DCR.0000000000002545
13. Al Abri R, Mathew J, Jeyaseelan L. Multiple mini-interview consistency and satisfactoriness for residency program recruitment: Oman evidence. *Oman Med J*. 2019;34(3):218-223. doi:10.5001/omj.2019.42
14. Altmaier EM, Smith WL, O'Halloran CM, Franken EA Jr. The predictive utility of behavior-based interviewing compared with traditional interviewing in the selection of radiology residents. *Invest Radiol*. 1992;27(5):385-389. doi:10.1097/00004424-199205000-00013
15. Asaad M, Elmorsi R, Ferry AM, Rajesh A, Maricevich RS. The experience of virtual interviews in resident selection: a survey of program directors in surgery. *J Surg Res*. 2022;270:208-213. doi:10.1016/j.jss.2021.09.011
16. Beran B, Kaljo K, Narayan R, Lemen P. An analysis of obstetrics-gynecology residency interview methods in a single institution. *J Surg Educ*. 2019;76(2):414-419. doi:10.1016/j.jsurg.2018.09.001
17. Burgos LM, Lima AADE, Parodi J, et al. Reliability and acceptability of the multiple mini-interview for selection of residents in cardiology. *J Adv Med Educ Prof*. 2020;8(1):25-31. doi:10.30476/jamp.2019.83903.1116
18. Camp CL, Sousa PL, Hanssen AD, et al. Orthopedic surgery applicants: what they want in an interview and how they are influenced by post-interview contact. *J Surg Educ*. 2016;73(4):709-714. doi:10.1016/j.jsurg.2016.03.009
19. Campagna-Vaillancourt M, Manoukian J, Razack S, Nguyen LH. Acceptability and reliability of multiple mini interviews for admission to otolaryngology residency. *Laryngoscope*. 2014;124(1):91-96. doi:10.1002/lary.24122
20. Carpinito GP, Khouri RK, Kenigsberg AP, et al. The virtual urology residency match process: moving beyond the pandemic. *Urology*. 2021;158:33-38. doi:10.1016/j.urology.2021.06.038
21. D'Angelo AD, D'Angelo JD, Beaty JS, et al. Virtual interviews—utilizing technological affordances as a predictor of applicant confidence. *Am J Surg*. 2021;222(6):1085-1092. doi:10.1016/j.amjsurg.2021.10.003
22. Dore KL, Kreuger S, Ladhani M, et al. The reliability and acceptability of the multiple mini-interview as a selection instrument for postgraduate admissions. *Acad Med*. 2010;85(suppl 10):60-63. doi:10.1097/ACM.0b013e3181ed442b
23. Finlayson HC, Townson AF. Resident selection for a physical medicine and rehabilitation program: feasibility and reliability of the multiple mini-interview. *Am J Phys Med Rehabil*. 2011;90(4):330-335. doi:10.1097/PHM.0b013e31820f9677
24. Gardner AK, D'Onofrio BC, Dunkin BJ. Can we get faculty interviewers on the same page? An examination

of a structured interview course for surgeons. *J Surg Educ*. 2018;75(1):72-77. doi:10.1016/j.jsurg.2017.06.006
25. Gilbart MK, Cusimano MD, Regehr G. Evaluating surgical resident selection procedures. *Am J Surg*. 2001; 181(3):221-225. doi:10.1016/S0002-9610(01)00550-5
26. Gordon EKB, Clapp JT, Heins SJ, et al. The role of the interview in residency selection: a mixed-methods study. *Med Educ*. 2020;54(11):1029-1039. doi:10.1111/medu.14248
27. Greene HL, Beattie HM, Russon AR, Johnson A, Stickley W, Goldberg R. Impact of a residency program information videotape on resident selection interviewing as a time-saving strategy. *J Med Educ*. 1985;60(4): 335-337. doi:10.1097/00001888-198504000-00010
28. Grober ED, Matsumoto ED, Jewett MA, Chin JL. The Canadian Urology Fair: a model for minimizing the financial and academic costs of the residency selection process. *Can J Surg*. 2003;46(6):458-462.
29. Hamel P, Boisjoly H, Corriveau C, et al. Using the CanMEDS roles when interviewing for an ophthalmology residency program. *Can J Ophthalmol*. 2007;42(2):299-304.
30. Hauge LS, Stroessner SJ, Chowdhry S, Wool NL. Evaluating resident candidates: does closed file review impact faculty ratings? *Am J Surg*. 2007;193(6): 761-765. doi:10.1016/j.amjsurg.2007.01.020
31. Hern HG Jr, Wills CP, Johnson B. Change to an informal interview dress code improves residency applicant perceptions. *West J Emerg Med*. 2015; 16(1):127-132. doi:10.5811/westjem.2014.11.22982
32. Isaksen JH, Hertel NT, Kjær NK. Semi-structured interview is a reliable and feasible tool for selection of doctors for general practice specialist training. *Dan Med J*. 2013;60(9):A4692.
33. Jacobs JC, Guralnick ML, Sandlow JI, et al. Senior medical student opinions regarding the ideal urology interview day. *J Surg Educ*. 2014;71(6):878-882. doi:10.1016/j.jsurg.2014.05.009
34. Kamboj AK, Chandrasekhara V, Simonetto DA, Raffals LE. How we did it: virtual interviews with an eye toward the future. *Am J Gastroenterol*. 2021;116(10): 1972-1975. doi:10.14309/ajg.0000000000001314
35. Kenigsberg AP, Khouri RK Jr, Kuprasertkul A, Wong D, Ganesan V, Lemack GE. Urology residency applications in the COVID-19 era. *Urology*. 2020;143:55-61. doi:10.1016/j.urology.2020.05.072
36. Lubarsky S, Young M. Assessing the multiple mini interview (MMI) for use in selecting applicants to a neurology residency program. *Neurology*. 2013; 80(suppl 7):P04.253. doi:10.1212/WNL.80.7_supplement.P04.253
37. Moran SK, Nguyen JK, Grimm LJ, et al. Should radiology residency interviews remain virtual? Results of a multi-institutional survey inform the debate. *Acad Radiol*. 2022;29(10):1595-1607. doi:10.1016/j.acra.2021.10.017
38. Ogunyemi D, Alexander C, Tangchitnob E, Kim DS. Mini surgical simulation, role play, and group and behavioral interviews in resident selection. *J Grad Med Educ*. 2016;8(3):410-416. doi:10.4300/JGME-D-15-00203.1
39. Rajesh A, Asaad M, Elmorsi R, Ferry AM, Maricevich RS. The virtual interview experience for Match 2021: a pilot survey of general surgery residency program directors. *Am Surg*. 2023;89(4):720-725. doi:10.1177/00031348211038555
40. Seifi A, Mirahmadizadeh A, Eslami V. Perception of medical students and residents about virtual interviews for residency applications in the United States. *PloS One*. 2020;15(8):e0238239. doi:10.1371/journal.pone.0238239
41. Shah SK, Arora S, Skipper B, Kalishman S, Timm TC, Smith AY. Randomized evaluation of a web based interview process for urology resident selection. *J Urol*. 2012;187(4):1380-1384. doi:10.1016/j.juro.2011.11.108
42. Shipper ES, Forrester J, Lau JN, Melcher ML. Evaluating the impact of blinded vs non-blinded interviews on the general surgery resident selection process. *J Am Coll Surg*. 2017;225(4 suppl 1):174-175. doi:10.1016/j.jamcollsurg.2017.07.396
43. Sklar MC, Eskander A, Dore K, Witterick IJ. Comparing the traditional and multiple mini interviews in the selection of post-graduate medical trainees. *Can Med Educ J*. 2015;6(2):e6-e13.
44. Swanson WS, Harris MC, Master C, Gallagher PR, Mauro AE, Ludwig S. The impact of the interview in pediatric residency selection. *Ambul Pediatr*. 2005;5(4):216-220. doi:10.1367/A04-149R1.1
45. Vadi MG, Malkin MR, Lenart J, Stier GR, Gatling JW, Applegate RL 2nd. Comparison of web-based and face-to-face interviews for application to an anesthesiology training program: a pilot study. *Int J Med Educ*. 2016;7:102-108. doi:10.5116/ijme.56e5.491a
46. VanOrder T, Robbins W, Zemper E. Residency program directors' interview methods and satisfaction with resident selection across multiple specialties. *J Am Osteopath Assoc*. 2017;117(4):226-232. doi:10.7556/jaoa.2017.040
47. Vermeulen MI, Kuyvenhoven MM, Zuithoff NP, van der Graaf Y, Damoiseaux RA. Dutch postgraduate GP selection procedure; reliability of interview assessments. *BMC Fam Pract*. 2013;14:43. doi:10.1186/1471-2296-14-43
48. Yamada T, Hiraoka E, Takemura Y. Predictive validity of past-behavioral versus situational questions in postgraduate medical admissions. *Acad Med*. 2021;96(2): 164-165. doi:10.1097/ACM.0000000000003821
49. Yamada T, Sato J, Yoshimura H, et al. Reliability and acceptability of six station multiple mini-interviews: past-behavioural versus situational questions in postgraduate medical admission. *BMC Med Educ*. 2017;17(1):57. doi:10.1186/s12909-017-0898-z

50. Yoshimura H, Kitazono H, Fujitani S, et al. Past-behavioural versus situational questions in a postgraduate admissions multiple mini-interview: a reliability and acceptability comparison. *BMC Med Educ*. 2015;15:75. doi:10.1186/s12909-015-0361-y
51. Association of American Medical Colleges. Interviews in GME: where do we go from here? Accessed December 31, 2023. https://www.aamc.org/about-us/mission-areas/medical-education/interviews-gme-where-do-we-go-here
52. Tiller D, O'Mara D, Rothnie I, Dunn S, Lee L, Roberts C. Internet-based multiple mini-interviews for candidate selection for graduate entry programmes. *Med Educ*. 2013;47(8):801-810. doi:10.1111/medu.12224
53. Kelly ME, Dowell J, Husbands A, et al. The fairness, predictive validity and acceptability of multiple mini interview in an internationally diverse student population—a mixed methods study. *BMC Med Educ*. 2014;14:267. doi:10.1186/s12909-014-0267-0
54. Kumar K, Roberts C, Rothnie I, du Fresne C, Walton M. Experiences of the multiple mini-interview: a qualitative analysis. *Med Educ*. 2009;43(4):360-367. doi:10.1111/j.1365-2923.2009.03291.x
55. Jerant A, Griffin E, Rainwater J, et al. Does applicant personality influence multiple mini-interview performance and medical school acceptance offers? *Acad Med*. 2012;87(9):1250-1259. doi:10.1097/ACM.0b013e31826102ad
56. Reiter HI, Lockyer J, Ziola B, Courneya CA, Eva K. Should efforts in favor of medical student diversity be focused during admissions or farther upstream? *Acad Med*. 2012;87(4):443-448. doi:10.1097/ACM.0b013e318248f7f3
57. Levine RB, Hebert RS, Wright SM. Resident research and scholarly activity in internal medicine residency training programs. *J Gen Intern Med*. 2005;20(2):155-159. doi:10.1111/j.1525-1497.2005.40270.x
58. Krumpal I. Determinants of social desirability bias in sensitive surveys: a literature review. *Qual Quant*. 2013;47(4):2025-2047. doi:10.1007/s11135-011-9640-9
59. Kelley K, Clark B, Brown V, Sitzia J. Good practice in the conduct and reporting of survey research. *Int J Qual Health Care*. 2003;15(3):261-266. doi:10.1093/intqhc/mzg031
60. Presser S, Couper MP, Lessler JT, et al. Methods for testing and evaluating survey questions. In: Presser S, Rothgeb JM, Couper MP, et al, eds. *Methods for Testing and Evaluating Survey Questionnaires*. Wiley & Sons; 2004:1-22.
61. Rudiawarni FA, Narsa IM, Tjahjadi B. Are emotions exacerbating the recency bias?: An experimental study. *Int J Trade Glob Mark*. 2020;13(1):61-70. doi:10.1504/ijtgm.2020.104913

John C. Lin, ScB, is a Medical Student, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA; **Daniel J. Hu, AB,** is a Medical Student, the Warren Alpert Medical School of Brown University, Providence, Rhode Island, USA; **Ingrid U. Scott, MD, MPH,** is Professor of Ophthalmology and Public Health Sciences, Penn State College of Medicine, Hershey, Pennsylvania, USA; and **Paul B. Greenberg, MD, MPH,** is Professor of Surgery (Ophthalmology), the Warren Alpert Medical School of Brown University, and Associate Chief of Staff for Surgery, VA Providence Healthcare System, Providence, Rhode Island, USA.

Funding: The authors report no external funding source for this study.

Conflict of interest: Paul B. Greenberg, MD, MPH, is a former ex-officio member of the Accreditation Council for Graduate Medical Education Medically Underserved Areas and Population Advisory Group (MUA/P) in 2022-2023 and received accommodations for the MUA/P meeting. Dr Greenberg is also a council member for the National Board of Medical Examiners (NBME) and received accommodations and meals for the NBME annual meeting.

Disclaimer: The views expressed here are those of the authors and do not necessarily reflect the position or policy of the US Department of Veterans Affairs or the US government.

Corresponding author: Paul B. Greenberg MD, MPH, Brown University, Providence, Rhode Island, USA, paul_greenberg@brown.edu

Received February 10, 2023; revisions received September 11, 2023, and January 11, 2024; accepted January 30, 2024.