Can We Get Faculty Interviewers on the Same Page? An Examination of a Structured Interview Course for Surgeons

Aimee K. Gardner, PhD,*,†,1 Brenna C. D'Onofrio, BA,† and Brian J. Dunkin, MD†,‡

*Department of Surgery, School of Allied Health Sciences, Baylor College of Medicine, Houston, Texas; †SurgWise Consulting, Houston, Texas; and ‡Department of Surgery, Houston Methodist Hospital, Houston, Texas

INTRODUCTION: Guidance on how to train faculty to conduct structured interviews and implement them into current screening processes is lacking. The goal of this study is to describe a structured interview training program designed specifically for surgeons and examine its effectiveness.

METHODS: Faculty involved in advanced surgical fellowship interviews completed a 20-item knowledge assessment and video-based applicant interview ratings before taking a half-day course on conducting structured interviews. The course consisted of evidence-based strategies and methods for conducting structured interviews, asking questions, and rating applicants in a highly interactive format. After the course, faculty again completed the knowledge assessment and provided ratings for 3 video-based applicant interviews.

RESULTS: All faculty members ($N = 5$) responsible for selecting fellows in minimally invasive and bariatric surgery completed the training. Faculty had an average of 15.8 ± 9.12 years in practice. Average performance on the precourse knowledge assessment was $35\% \pm 6.12\%$ and the group was unable to achieve acceptable agreement for applicant interview scores for any of the competencies assessed. After the course, faculty demonstrated significant improvements ($p < 0.01$) on the knowledge assessment, more than doubling their scores on the pretest with average scores of $80\% \pm 9.35\%$. Faculty also improved their interrater agreement of applicant competency, with 80% of the applicant interview ratings within 2 points of each other.

CONCLUSIONS: Implementation of a half-day course designed to teach principles and skills around structured interviewing and assessment demonstrated significant improvements in both interviewing knowledge and inter-rater agreement. These findings support the time and resources required to develop and implement a structured interview training program for surgeons for the postgraduate admissions process. (J Surg Ed ∎:∎∎∎-∎∎∎. © 2017 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: interview, selection, residency

COMPETENCIES: Systems-Based Practice

INTRODUCTION

Program directors have seen a steady increase in the applicant pool for general surgery residency positions over the past 5 years,[1] creating a significant burden on programs and administrators to identify applicants that best fit into their training environment.[2] The screening process for filling residency positions typically involves review of an applicant's United States Medical Licensing Examination step 1 scores, grades in third year of medical school, and letters of recommendation. Those who pass a predefined threshold are invited for an on-site interview.[3] This last hurdle, the interview, has been identified as the most important factor in determining final selections.[4-6]

Despite the substantial weight given to the interview in residency selection,[4-7] studies have documented the subjectivity of these interviews by noting extreme amounts of variability in duration, structure, and methods used both within and between programs.[3] Not only does this lack of structure result in low interrater reliability[8] and inability to

Correspondence: Inquiries to Aimee K. Gardner, PhD, School of Allied Health Sciences, Baylor College of Medicine, MS: BCM115, DeBakey Building, M108K, One Baylor Plaza, Houston, TX 77030; fax: (713)-798-7694; e-mail: aimee.gardner@bcm.edu

¹ Twitter: @AimeeGthePHD.

TABLE 1. Course Components and Delivery Methods

Course Topic	Didactics	Video or Audio-Based Case Review	Small Group Discussion	Role Play	Other Active Learning Strategies
Part I: Background					
Structured interview basics	x		x		
Question development	x		x		
Biases in interviews		x	x		
Part II: Asking Questions					
Getting complete responses		x	x		
Types of interviewing questions	x		x		
Unacceptable and illegal questions		x	x	x	x
Taking notes		x			x
Part III: Assessment					
Assigning ratings		x	x	x	x
Motivational fit		x	x		
Integrating data	x		x		
Part IV: Putting It All Together					
Interview day basics	x				
Review		x	x		

predict future performance,[2] but it also contributes to a high incidence of potentially illegal questions asked during surgery residency interviews.[8] For these reasons, educators have criticized the traditional resident selection interview for having "dubious value."[2]

Structured interviews, in contrast, rely on more objective evaluation methods, such as consistently asking only questions related to position requirements, providing training on interviewing skills, and rating interviewees using established scoring formats.[9] Numerous meta-analyses have shown the ability of structured interviews to predict performance across a wide array of settings, at a rate twice as high as unstructured interviews.[10-12] Further comparison of these techniques has revealed that it requires a minimum of 4 unstructured interviews to achieve the levels of reliability and predictive validity that one interviewer using structured interview techniques would attain.[13] For these reasons, incorporation of structured interviews has been considered as a best practice among medical educators.[2,14]

Unfortunately, guidance on how to train faculty to conduct these interviews and implement into current screening processes is lacking. The goal of this study is to describe a structured interview training program designed specifically for surgeons and examine its effectiveness.

METHODS

Pretraining Assessment

Before beginning the course, participants completed 2 forms of pretraining assessments. The first consisted of a 20-item knowledge assessment, which examined participant's baseline knowledge regarding characteristics of structured interviews, biases in interviewing, unacceptable questions, question formats, note taking, and rating techniques. Participants then reviewed a video depicting an interview between an applicant (Nico) and faculty surgeon, and rated the applicant's interpersonal skills, communication, and adaptability according to a behavioral anchor rating scale specific to the respective competencies (1 = much less than acceptable; 5 = acceptable; 10 = much more than acceptable). Participants were asked not to discuss the video or their evaluations.

Curriculum

Course components were led by 2 instructors (A.K.G. and B.C.D.) and are detailed in Table 1. The course is divided into 4 overarching themes as follows: background, asking questions, assessment, and putting it all together, with evidence-based strategies and methods highlighted throughout. In part I participants were provided with an introduction to structured interviews, including discussion of essential characteristics, comparisons with traditional/ unstructured interviews, validity evidence supporting their use, and data regarding applicant perceptions. The course then provided an overview of how questions are developed in structured interviews and common biases prevalent among interviewers.

Part II consisted of topics pertaining to asking questions in interviews, including techniques to obtain complete responses from applicants, types of questions and when to use them, unacceptable and illegal questions (e.g., asking

TABLE 2. Percentage of Agreement Among Interviewers

Interviewee	Competency Assessed	% Agreement (%)
Nico (pre)	Adaptability	0
	Communication skills	0
	Interpersonal skills	0
Training		
Nico (post)	Adaptability	100
	Communication skills	100
	Interpersonal skills	0
Silvia (post)	Conflict management	100
	Communication skills	100
	Interpersonal skills	0
Taylor (post)	Problem solving	100
	Stress management	100
	Communication skills	100
	Interpersonal skills	100

Note: Competency ratings were rated on a scale of 1-10 with varying anchors. Agreement = all faculty members were within 2 points of each other.

about family planning, marital status, and demographics), and note taking.

Part III focused on assessments, including an introduction to behavioral anchor rating scales, gauging motivational fit, and integrating data among multiple interviewers. In this portion of the training program, participants also engaged in frame of reference training[15,16] with multiple iterations of assessment and consensus discussion based on to ensure that they were all using the rating scales in the same manner. Audio examples reflecting a variety of competencies were used to prompt these ratings and discussions.

Finally, in part IV participants were provided with an overview of tips and tricks for the interview day, including preparing for the interview, opening the interview, gathering information, closing the interview, and making evaluations. The course took 5 hours to complete and used various delivery methods, including didactics, video and audio-based case review, small group discussion, role play, and other active learning strategies.

Posttraining Assessment

At the conclusion of the training session, participants observed 3 video-based interviews and completed assessments on the hypothetical "applicants," Nico, Silvia, and Taylor who were applying to a surgery program and were asked questions about adaptability (Nico), conflict management (Silvia), and multitasking and stress management (Taylor). Participants rated these videos according to the same scoring scale (1 = much less than acceptable; 5 = acceptable; 10 = much more than acceptable) used in the pretraining assessments. Afterward, they completed the same 20-item knowledge assessment used in the pretest, along with a 14-item course evaluation that assessed course delivery aspects (i.e., relevance, amount of material, pace, instructor quality, and opportunities to apply new material)

and specific content areas (structured interview basics, biases in interviews, getting complete responses, unacceptable questions, taking notes, assigning ratings, motivational fit, and putting it all together) using a 1 (critically deficient) to 5 (exceeds expectations) scale.

Data Analysis

Basic descriptives and analyses were obtained with SPSS version 24.0 (IBM; Chicago) Pretraining and posttraining changes on the assessments were performed with paired samples t tests. Interrater agreement was calculated by the agreement of ratings (i.e., 2 or fewer points away on the 1-10 scale) for each competency as has been previously reported in the medical education literature.[17] Other estimates of interrater agreement (i.e., intraclass correlations) require multiple items per competency and rely on larger sample sizes.

RESULTS

All faculty members ($N = 5$) responsible for selecting fellows in a minimally invasive and bariatric surgery program completed the training. Faculty averaged nearly 16 years (15.8 ± 9.12 years) in practice at academic institutions in which they have been responsible for conducting interviews for medical students, residents, and fellows.

Before the course, the average performance on the precourse knowledge assessment was 35% ± 6.12%. Topic areas that received the lowest scores were interviewing techniques and unacceptable questions. Faculty also demonstrated wide variability on the precourse ratings of applicant competency. As shown in Table 2, faculty did not provide ratings within 2 points of another for any of the competencies assessed, including adaptability, communication, and interpersonal skills.

After participating in the course, faculty demonstrated significant improvement on the knowledge assessment, more than doubling their scores on the pretest with average postcourse scores of 80% ± 9.35% ($p < 0.01$ compared to pretest). Post hoc analyses revealed that improvement by topic included structured interview basics (50% → 80%, $p < 0.01$), interviewing techniques (33% → 89%, $p < 0.01$), assigning ratings (40% → 87%, $p < 0.05$), and unacceptable questions (20% → 65%, $p < 0.05$). These data are shown in the Figure.

Table 2 illustrates that faculty also improved on their interrater agreement of applicant competency, with 80% (8/10) of the ratings within 2 points of each other. The competency that did not receive uniform agreement on 2 of the 3 video-based interviews was interpersonal skills.

Participants rated the course highly, with an overall course evaluation of 4.54 ± 0.53. Average scores for the course delivery were 4.51 ± 0.30 and 4.55 ± 0.76 for course content.
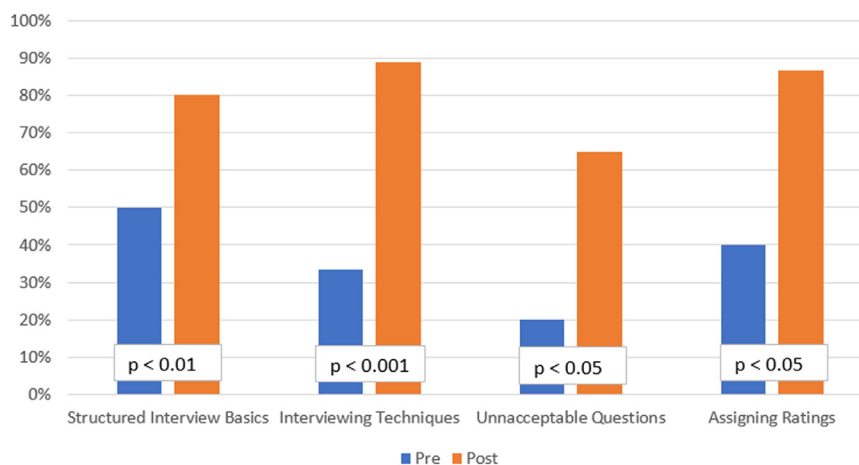
FIGURE. Knowledge assessment scores by topic.

DISCUSSION

In 1979, a legislative act required the University of Texas Medical School at Houston to admit 50 additional applicants late in the admissions season. These 50 applicants were selected from a pool that was initially rejected based largely on impressions from unstructured interviews in which interviewers asked questions regarding any content area in whatever way they chose. This gave the medical school a unique opportunity to evaluate the value of basing selection decisions on an unstructured interview process. At the end of the first year of postgraduate training for this expanded class, there were no meaningful differences in preclinical evaluations, clinical performance, honors earned, or attrition rates between the initially accepted and initially rejected groups.[18] Other large-scale field studies[19,20] have come to the same conclusion—that unstructured interviews provide little validity evidence for screening applicants.

The goal of our study was to describe and evaluate the components of a structured interview training program designed specifically for surgeons. Through case-based review, small group discussion, role play, and various active learning activities, the program was able to more than double the preexisting knowledge-base of surgery faculty with an average of 16 years of experience conducting unstructured interviews. Additionally, faculty more than tripled their preexisting knowledge-base in differentiating acceptable from unacceptable questions. The practical significance of this finding cannot be understated. Given that applicants to surgery programs are 3 times more likely than nonsurgical applicants to be asked at least one potentially illegal question during their interview,[8] this improvement represents a key shift in interviewer awareness of what inappropriate questions are and how to avoid asking them.

Our findings also demonstrate that the course significantly improved faculty agreement in assessing an applicant's responses to interview questions. Before the course, this group of faculty, who has been conducting interviews together for the past 4 years, was in agreement 0% of the time. Through multiple iterations of practice and discussion using powerful tools like frame of reference training, this agreement rose to 80%. These changes reflect one of the key characteristics and benefits of structured interviews—interviewers are using the same rating tools in the same way. Overall, these data suggest that not only do these course participants now better understand the components of structured interviews and how to conduct them, but they are also better equipped with skills to rate applicants during the interview process.

The implications of these findings should be considered in light of the substantial costs associated with conducting residency interviews. A recent survey of plastic surgery applicants found that students spend up to $10,000 traveling to interviews.[21] Estimates suggest that residency programs spend about the same amount for recruiting just one first-year resident.[22] National surveys of program directors in other specialties have found that the median total cost of recruitment per program is about $148,000.[22] Considering this considerable investment of personnel, time, and resources, it is critical that interview sessions be conducted efficiently and in a manner that maximizes their reliability and validity. Structuring the interview process has been shown to improve reliability, validity, fairness, and applicant perceptions.[23,24] It is also efficient in that 1 structured interview is as effective as 4 unstructured ones.[13] Thus, programs hoping to optimize the efficiency and effectiveness of their selection processes should train their faculty on how to conduct structured interviews.

Despite the importance of these findings, the study does have some limitations. It was conducted at a single institution with a small group of surgeons who represent teaching faculty for an advanced surgery fellowship program. Although this small group was ideal for the interactive training exercises and discussions, and represents the entire cohort of decision makers for fellowship selection, future

research should investigate the efficacy of this curriculum for larger programs with more faculty members. Furthermore, this study does not investigate the extent to which the knowledge and skills learned from the course were retained. We were able to show immediate knowledge gains and agreement among faculty, but a longitudinal study design is needed to fully understand the effect of this curriculum on knowledge retention. Development of "refresher" training programs may be needed on a recurring basis. Additionally, this work did not measure if these skills were applied to actual interviews, although there are plans to obtain this transfer of training data as well. Finally, the ultimate success of a rigorous assessment program during selection is success of those who are hired. Although we have demonstrated that we were able to get faculty on the same page, examination of the effect of using comprehensive and structured competency-based interview questions is needed.

CONCLUSION

This study demonstrated that experienced surgery faculty members who have been conducting selection interviews for many years have a low baseline understanding of evidence-based interview techniques and very low interrater reliability when assessing applicant interview responses. Implementation of a well-designed half-day course focused on teaching principles and skills for conducting structured interviews along with rehearsing the use of powerful assessment tools, demonstrated significant improvement both in interviewing knowledge and interrater agreement. These findings support the time and resources required to develop and implement a structured interview training program for surgeons for the postgraduate admissions process.

REFERENCES

1. Results of the Match. The National Resident Matching Program. Available at: ⟨http://www.nrmp.org/wp-content/uploads/2015/05/Main-Match-Results-and-Data-2015_final.pdf⟩. Accessed on February 13, 2017.

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.

3. Makdisi G, Takeuchi T, Rodriguez J, Rucinski J, Wise L. How we select our residents—a survey of selection criteria in general surgery residents. *J Surg Educ*. 2011;68(1):67-72.

4. Smilen S, Funai E, Bianco A. Residency selection: should interviewers be given applicants' board scores? *Am J Obstet Gynecol*. 2001;184(4):508-513.

5. Provan JL, Cuttres L. Preferences of program directors for evaluation of candidates for post graduate training. *CMAJ*. 1995;153(7):919-923.

6. Wagoner NE, Suriano R. Program directors' responses to a survey on variables used to select residents in a time of change. *Acad Med*. 1999;74(1):51-58.

7. Prager JD, Myer CM, Hayes KM, Myer CM, Pensak ML. Improving methods for resident selection. *Laryngoscope*. 2010;120(12):2391-2398.

8. Hern HG, Trivedi T, Alter HJ, Wills CP. How prevalent are potentially illegal questions during residency interviews? A follow-up study of applicants to all specialties in the national resident matching program *Acad Med*. 2016;91(11):1546-1553.

9. Dipboye RL, Wooten K, Halverson SK. Behavioral and situational interviews. Thomas JC, editor. Comprehensive Handbook of Psychological Assessment, 4, Industrial and Organizational Assessment. Hoboken, NJ: John Wiley & Sons Inc, 2004. p. 297-316.

10. McDaniel MA, Whetzel DL, Schmidt FL, Maurer SD. The validity of employment interviews: a comprehensive review and meta-analysis. *J Appl Psychol*. 1994;79(4):599-616.

11. Huffcuff AI, Conway JM, Roth PL, Stone MJ. Identification and meta-analytic assessment of psychological constructs measured in employment interviews. *J Appl Psychol*. 2001;86(5):897-913.

12. Wiesner W, Cronshaw S. A meta-analytic investigation of the impact of interview format and degree of structure on the validity of the employment interview. *J Occup Psychol*. 1988;61:275-290.

13. Huffcutt AL, Arthur W. Hunter and Hunter (1984) revisited: interview validity for entry-level jobs. *J Appl Psychol*. 1994;79:184-190.

14. Bandiera G, Abrahams C, Ruetalo M, Hanson MD, Nickell L, Spadafora S. Identifying and promoting best practices in residency application and selection in a complex academic health network. *Acad Med*. 2015;90(12):1594-1601.

15. Day DV, Sulsky LM. Effects of frame-of-reference training and information configuration on memory organization and rating accuracy. *J Appl Psychol*. 1995;80(1):152-167.

16. Gardner AK, Russo MA, Jabbour II, Kosemund M, Scott DJ. Frame-of-reference training for simulation-based

intraoperative communication assessment. *Am J Surg.* 2016;212(3):548-551.

17. Patrick LE, Altmaier EM, Kuperman S, Ugolini K. A structured interview for medical school admission, phase 1: initial procedures and results. *Acad Med.* 2001;76(1):66-71.

18. DeVaul R, Jervey F, Chappell J, Caver P, Short B, O'Keefe S. Medical school performance of initially rejected students. *J Am Med Assoc.* 1987;257(1):47-51.

19. Millstein RM, Wilkinson L, Burrow GN, Kessen W. Admission decisions and performance during medical school. *J Med Educ.* 1981;56(2):77-82.

20. Carroll JS, Wiener RL, Coates D, Galegher J, Alibrio JJ. Evaluatoin, diagnosis, and prediction in parole decision making. *Law Soc Rev.* 1982;17(1):199-228.

21. Wood JS, David LR. Outcome analysis of factors impacting the plastic surgery match. *Ann Plast Surg.* 2010;64(6):770-774.

22. Brummond A, Sefcik S, Halvorsen AJ, et al. Resident recruitment costs: a national survey of internal medicine program directors. *Am J Med.* 2013;126(7): 646-653.

23. Campion MA, Palmer DK, Campion JE. A review of structure in the selection interview. *Person Psychol.* 1997;50:655-702.

24. Levashina J, Hartwell CJ, Morgeson FP, Campion MA. The structured employment interview: narrative and quantitative review of the research literature. *Person Psychol.* 2013;67(1):241-293.