Simulation-Based Selection of Surgical Trainees: Considerations, Challenges, and Opportunities

Aimee K Gardner, PhD, E Matthew Ritter, MD, FACS, John T Paige, MD, FACS, Rami A Ahmed, DO, MHPE, Gladys Fernandez, MD, Brian J Dunkin, MD, FACS

When faced with a large number of applicants for a limited number of positions, residency and fellowship programs in surgery must adopt some kind of selection procedure. For residency programs, these selection procedures are traditionally based on academic achievement, knowledge of science-related subjects, and cognitive abilities.[1] Although these cognitively oriented variables have been shown to be good predictors of academic performance in medical training,[2] educators still struggle to implement effective and efficient ways to identify individuals who will be successful in their training programs.[1] For example, many suggest that screening for decision-making skills, emotional intelligence, or even innate dexterity might be helpful for various specialties.[2,3] Although some current screening methods, such as letters of recommendations or interviews, can variably capture some of these competencies, few standardized assessments exist. As these constructs are hard to assess with traditional paper-and-pencil formats, innovative screening and assessment programs, including the use of simulation, might be needed. For example, placing an applicant in a realistic scenario in which he or she might have to demonstrate problem solving, interpersonal, and/or leadership skills can provide unique information that decision makers might not otherwise obtain using solely cognitive assessments.

It is possible that simulation can provide decision makers with important information about applicant suitability, but little is currently known about the feasibility and use of incorporating simulation exercises into the screening and selection process in medical education. A better understanding of if and how simulation can be used to help inform selection decisions among surgical educators is needed. What follows is a summary of these discussions, with an overview of the strengths and limitations of the use of simulation in the selection of candidates into training programs.

VALUE OF SIMULATION-BASED SELECTION

Supporters of using simulation-based selection (SBS) for applicants point to the notion of behavioral consistency,[4] which posits that the behavior of candidates in situations similar to those encountered in the hospital will provide good predictions of actual behavior in the clinical setting. Placing applicants in situations that will be experienced later during training provides a "realistic preview" of how that candidate might perform in a training program. Importantly, these simulations can take a variety of forms, as shown in Table 1.[5] They can be situational judgment tests (SJTs) in which applicants are presented with situations that they will likely encounter during training and asks candidates to respond in one of two ways, what they would do or what they should do, given the situation. Or, SBS can take the form of work samples, in which candidates are asked to perform hands-on tasks (eg a skill or procedure) that are physically and/or psychologically similar to those performed in training. Finally, SBS can be more high-fidelity "assessment centers" (eg role plays) meant to measure a wide array of nontechnical competencies, such as interpersonal skills, communication skills, organizing, judgment, and analytical skills. Regardless of form, the sole purpose of SBS remains the same: to make decisions based on data derived from applicant performance when completing a task, interacting with others, or working with systems. As will be discussed, SBS has numerous benefits over traditional selection processes, including allowing flexibility in implementation, capturing a wider array of candidate competencies, "test driving" the applicant, potentially enhancing the validity of selection decisions, and providing a realistic preview to applicants.

Disclosure Information: Nothing to disclose.

Disclosures outside the scope of this work: Dr Ahmed receives support as a paid consultant to NASA and Simulab and is also supported by royalties from Simulab.

Presented at the 16th International Meeting on Simulation in Healthcare, San Diego, CA, January 2016.

Received April 23, 2016; Revised May 17, 2016; Accepted May 23, 2016.
From the Department of Surgery, University of Texas Southwestern Medical Center, Dallas (Gardner), Department of Surgery, Houston Methodist Hospital, Houston (Dunkin), TX, Department of Surgery, Uniformed Services University and Walter Reed National Military Medical Center, Bethesda, MD (Ritter), Department of Surgery, Louisiana State University Health Sciences Center, New Orleans, LA (Paige), Department of Emergency Medicine, Summa Akron City Hospital, Akron, OH (Ahmed), and Department of Surgery, Baystate Medical Center, Springfield, MA (Fernandez).

Correspondence address: Aimee K Gardner, PhD, Department of Surgery, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390. email: aimee.gardner@utsouthwestern.edu

http://dx.doi.org/10.1016/j.jamcollsurg.2016.05.021
ISSN 1072-7515/16

Table 1. Overview of Situational Judgment Tests, Work Samples, and Assessment Centers

Method	Definition	Example	Advantages	Disadvantages
Situational judgment tests	Applicants are presented with a description of a work-related scenario and asked to exercise their judgment by choosing alternative courses of action given the situation.	You are a junior resident rotating on a service that frequently interacts with EM residents and attendings. Unfortunately, the EM department and surgery house staff have a history of confrontation. As a result, you find that your actions are constantly being scrutinized and questioned by EM faculty and residents. Your attending has received multiple complaints about your interpersonal behavior, although you are certain they are unfounded. Which of the following actions should you take, from most to least appropriate? A. Tell your attending the complaints are without merit. B. Talk to your colleagues to see if they are having similar experiences on this rotation. C. Speak with the EM faculty to inquire more about how your behavior is being perceived. D. Apologize to EM faculty and residents and monitor your behavior closely. E. Do nothing, and keep to yourself until the rotation is over. F. Speak with the program director about these issues.	Easily administered and scored, especially when using video-based or digital technology to administer and record answers; applicants can see the link between SJT scenarios and the work. As a result, they take the test seriously and try to do well. Also, the test provides a good preview of what the job will be like; SJTs are generally not susceptible to "practice effects" (ie improved performance on the assessment if the applicants complete it more than once). It is harder for dishonest applicants to remember and disclose the longer SJT questions to other applicants.	SJTs help measure applicants' responses to ambiguous conditions. Good SJT scenarios contain a rich set of details, only some of which help determine the correct answer. If the test provides insufficient detail, the question does not fully test an applicant's ability to discern the relevant information and respond accordingly; if correct answers are too obvious, the SJT can become a test of what applicants know they "should" do on the job rather than what they would actually do.
Work samples	Evaluates applicants' job-related skills by having them perform actual activities or tasks that are physically or mentally similar to the duties they would perform on the job.	A work sample for a general surgery residency position might involve having applicants perform knot-tying or suturing tasks.	They generally have high predictive validity; applicants are less able to "fake" proficiency; Applicants view them as fair because they can see the relationship to the job; work sample tests provide applicants with a job preview to better inform their decision on whether they are a good fit for the job.	They generally do not measure aptitude or future potential; their scope is limited to only the competencies needed for the specific activity carried out during the test; they are not very useful for tasks that take a long time to complete.
Assessment centers	Uses multiple techniques and multiple assessors to produce judgments about the extent to which a participant displays selected competencies.	Assessment centers always use more than one exercise to measure the different applicant dimensions under review. Some of the more common exercises used in assessment centers include in-basket exercises (in which an applicant manages a set of tasks provided in a simulated in basket), leaderless group discussions, structured interviews, and oral presentations.	They have moderately high validity ratings, meaning that they have been found to be good predictors of job performance, especially in terms of leadership abilities; applicants view them as fair because they can see the relationship between the exercises and the job; assessment centers provide applicants with a job preview to better inform their decision on whether they are a good fit for the job.	The key disadvantage to assessment centers is that they are resource intensive. They take time and expertise to develop and organize. They require multiple, trained raters. They require space, equipment, and materials to administer. All of these resources amount to a fairly significant cost.

Adapted from the US Merit Systems Protection Board.[5]
EM, emergency medicine; SJT, situational judgment tests.

Flexibility

Programs looking to add simulation to the selection process have a wider range of design and delivery choices than with traditional assessment methods.[6] Simulation can capture as many customized competencies as a program wishes. For example, fellowship programs characterized by high-paced decision making (eg critical care, trauma) might wish to place candidates in simulations in which they must demonstrate good judgment and leadership in time-constrained situations. In contrast, subspecialties in which trainees often have to interact with sensitive populations (ie pediatrics, palliative care) might wish to focus simulation assessment on emotional intelligence, communication, and interpersonal skills. Specialties with a heavy focus on technical skills (eg surgery, interventional radiology) might wish to focus on procedural aptitude. General surgery residency programs can choose to use a combination of all of these to ensure a basic level of all competencies needed to be successful in a program. Additionally, simulations can take a variety of formats. They can occur in a paper-and-pencil format or be computer-based. Candidates can interact with a single individual or with a team. They can be instructed to act casually (eg teaching a new skill) or in highly structured ways (eg selecting a behavior from a list of alternative responses). Additionally, behavior can be assessed immediately or recorded for later evaluation. Simulation-based selection systems can be created to meet the specific needs of a program and/or specialty, and be adapted to complement available resources.

Assessing a broader array of competencies

As noted previously, the flexibility of simulation allows decision makers to assess a broad array of applicant knowledge, skills, and abilities. Scholars are increasingly noting that medical knowledge and technical skill are not enough to successfully perform in training programs.[7] For example, when interviewing a sample of general surgery program directors, Sanfey and colleagues[8] noted that none of those interviewed had ever terminated a resident or denied a resident promotion because of poor technical skills. Rather, it is critical that physicians possess high levels of interpersonal, communication, leadership, teamwork, and problem-solving skills. Unfortunately, the competencies mentioned are difficult to capture in a standardized fashion using current selection systems that only consider standardized test scores, letters of recommendation, and unstructured interviews.[9] Overlooking other key competencies might be why many programs experience such high levels of remediation and attrition. A 20-year study of Yale residents by Longo[10] revealed that 30% of categorical general surgery residents at their institution failed to complete the general surgery program. Of note, the majority of these trainees (77%) left voluntarily or involuntarily before beginning their third year. The authors concluded that "although programs recruit residents primarily on cognitive factors, such as grades on clinical rotations and standardized board scores, it is the vital noncognitive issues that are associated with failure." It is clear that when placed in the actual demands of a clinical training environment, some individuals are unable or unwilling to meet those demands. By re-creating these environments with simulated exercises, programs might be able to proactively identify these individuals and make selection decisions accordingly. At a minimum, placing candidates in a situation in which they must learn a new skill can provide important insight into their "learning aptitude" or ability to receive feedback. Similarly, SBS might also allow candidates to gain insight about the teaching ability and environment in which they would be immersed if selected. Of course, to be maximally useful in practice, more research is needed to better understand which competencies are best measured and by which types of simulations.

Previewing the applicant

As mentioned previously, simulation allows programs to place candidates in situations that are meant to reflect the demands of clinical training, providing some foresight as to how an applicant would perform in the clinical environment. Each simulation should represent an important challenging aspect of the program's work environment that realistically reflects the demands, values, and/or culture of the program and allows for a standardized evaluation of an applicant's performance. In this way, simulations would parallel the notion of Sabermetrics, the subject of the movie *Moneyball*.[11] Sabermetrics originated in baseball and uses predictive modeling based on specific player statistics. Of note, making decisions on these pieces of data has been found to be more effective than talent scouts with years of wisdom and experience in selecting players who win games.[12] In the same way, simulation might allow for the opportunity to observe applicants perform in a more comprehensive sport and gather data on their performance rather than limiting assessment to candidate's knowledge about the sport. If put into place, surgical educators can similarly find that data provided from these simulations allow decision makers to make more successful judgments about candidates. Training programs might be wise to consider the value of watching "players play" before putting them "on the team" and "in the game."

Validity

Although there is a paucity of research investigating the validity of using SBS for medical trainees in the United

States, there has been some work performed on this topic in the United Kingdom. Specifically, Lievens and Patterson[3] worked with 20 physician experts during the course of 2 days to create 186 SJTs to assess communication, empathy, professional integrity, coping with pressure, and problem solving. Through pilot testing and psychometric analyses, the authors were able to finalize a set of 50 items that applicants could complete in 90 minutes. Their work demonstrated that performance on SJTs and a variety of simulation exercises used for selection purposes predicted later performance among 196 general practitioners in training. Importantly, their work also noted that these SJTs and simulations provided incremental validity above and beyond knowledge tests. As a result, many of these items remain as part of the advanced-level high-stakes exam in the United Kingdom. Other findings have shown that video-based SJTs created under similar methodologies to capture interpersonal skills are able to predict medical school performance over cognitive ability exams among a sample of 610 students in Belgium.[1] Because of these relationships and the ability of the video-based SJT to reduce adverse impact, this exam is included in the Flemish Admission Exam for Medical and Dental Studies. There is also a strong foundation of SBS literature from other domains. For example, work samples have been found to be among the most valid tools for predicting employee performance across a wide range of settings and occupations, from entry level to executive leadership positions.[13] These findings are likely a manifestation of the fact that simulations are performance-based assessments and are less vulnerable to answer manipulation toward perceived social desirability that can occur in personality inventories and interviews.[6] Additionally, during simulations, the candidate behaviorally demonstrates (within a representative situational context) the underlying capacities (ie technical skill, earlier learning, personality traits, decision making, etc) that other selection processes cannot measure in a nonbehavioral format. When designed properly, simulations should serve as a miniature replica of the role and environment for which a person is applying. In addition, beyond actual empirical validity for selection decisions, simulations have the advantage of looking valid, given their high degree of fidelity and similarity to the clinical environment.

Realistic preview to applicants

One of the most important characteristics of SBS is that it provides applicants with a preview of what it would be like to be in a particular program.[14] Only during a simulation can applicants be confronted with vividly representative real-life challenges that will occur during training and mimic feelings of pressure associated with a particular training position. By being placed in context-specific scenarios, applicants can better understand the unique demands they might face in a program and gauge their opinions of it. This realistic preview can help candidates evaluate their fit with the program. Importantly, this realistic preview also allows undesirable or uninterested candidates to self-select out of the program, saving both time and money. As a result, realistic job previews have been shown to lead to lower turnover.[15] Additionally, they increase engagement among applicants. In fact, meta-analyses have indicated that simulations are consistently perceived as one of the most favorable by candidates compared with all other selection techniques.[16] As programs move closer to achieving a 100% overlap between the simulation content and the target job, programs will be able to provide a more realistic job tryout in a safe environment. Simulation-based selection allows programs to create a realistic experience for applicants and also capitalize on the rich data obtained from simulation exercises to predict success in a training program.

LIMITATIONS OF SIMULATION-BASED SELECTION

Although the use of SBS might prove valuable in some industries, there is little research so far to suggest that these can benefit complex medical education systems. Not only are the training candidates themselves an extremely unique subset of the population who have already demonstrated their aptitude and abilities by making it through the rigorous medical training prerequisites and/or preliminary training programs, but the training environment is a constantly diverse and evolving setting. The demands of a physician are extremely broad, and no one test, exercise, or exam can comprehensively account for an individual's aptitude and future success. In addition, it is unclear whether a better selection system is even needed. The following is a summary of limitations of implementing SBS into surgical education programs, organized by discussions of validity and utility, applicant reactions, mistaking trainees for employees, and resources.

Validity and utility

Perhaps one of the largest concerns about using SBS is the unknown validity and utility of applying this process to the selection of medical trainees. Although there might be data from industry supporting the concept of using simulation exercises for high-volume selection, there are unique characteristics of medical education that warrant consideration. Residency and fellowship programs already have access to years of work sample information, which is reflected

in clinical evaluations and faculty recommendations. The primary issue, then, is how to capitalize on that information. Efforts might be more efficacious if they were directed toward using the plethora of information that is already accessible, rather than developing entirely new systems to re-create the crux of those data. Unfortunately, those who have attempted to organize and scrutinize the predictive validity of the information gathered in the current system have been left wanting for more useable data.[17] For example, reviews have demonstrated that our current system, based heavily on unstructured letters of recommendation and interviews, are unreliable in predicting resident success.[18-20]

Additionally, more work is needed to convincingly demonstrate the benefit of incorporating simulation into the selection process. Research supporting work samples in manufacturing or corporate environments might not be the best parallel for medical trainees. Physicians in training are not evaluated by sales figures or number of widgets constructed. Instead, their success is evaluated by ability to train, adapt, and perform in highly stressful and dynamic situations. Research from other domains might be more relevant. For example, the National Football League (NFL) spends substantial time, money, and resources to predict which college quarterbacks will be drafted into the NFL and have successful professional careers. Unfortunately, researchers have noted that college and combine statistics for drafted quarterbacks are not reliably associated with, or predictive of, success in the NFL.[21,22] Authors have concluded that it seems unlikely that collecting even more statistics on performance of college quarterbacks will help prediction models, as they already incorporate a plethora of quantitative measures. Instead, they assert that there are complex factors that are inherently unmeasurable and/or random that play a major role in determining who will be successful performer.[22] Similar results have been found for selection of running backs and wide receivers. In fact, correlation of measures taken at the NFL combine with multiple measures of success in the NFL resulted in fewer significant relationships than would have been predicted by chance alone.[23] After reviewing the predictive ability of the SBS process known as the NFL combine, authors concluded that "NFL team owners and managers must use caution in determining the value of the combine in drafting players."[23] The same statement can be applied to the selection of medical trainees; there are likely too many variables to assess and too many intricate factors that contribute to a successful medical trainee that an SBS system can comprehensively and accurately capture. Until these complexities are more clearly understood, simulation has little value for this purpose.

Applicant reactions

Proponents of SBS suggest that applicants might enjoy simulation exercises, be more engaged in the application process, and perhaps view a program as being more "cutting edge" if they include simulations in their selection process. However, it must be acknowledged that this process is already incredibly stressful for residency and fellowship candidates who must juggle numerous interviews and fly across the country or continents. Adding obstacles to the current system, such as requiring them to demonstrate a wide variety of competencies in a high-pressure situation and participate in more lengthy visits, will only increase frustration for applicants. Challenging and testing future trainees on initial introduction might not create the supportive image that programs need to attract high-quality applicants. Opponents of SBS might only get on board if proposed changes in selection strategies had high-level (ie national boards and accrediting agencies) support and uniform adoption to not isolate individual programs. Although those who are concerned about dwindling numbers of practitioners in surgery[24] might believe that SBS would increase the risk of turning away applicants, there are few data to suggest that few (if any) surgery programs are unable to match categorical positions. For example, the National Resident Matching Program data show that only 2 of the 1,224 categorical general surgery PGY1 positions in the United States were unfilled in 2015.[25] These data also show that the applicant pool for the specialty at large has been increasing steadily during the past 5 years, and that unmatched spots represent only 0.003% of total positions. Although programs have legitimate concerns in ensuring that the on-site visit to sell the program to each applicant, the increasing applicant pool suggests that now might be a good time to implement new processes.

Trainees vs employees

Although proponents will point to a number of industries outside of medical education that use these selection practices, this application fails to acknowledge that other industries are hiring individuals who must already have a certain level of knowledge, skills, or abilities to complete the assessment. For positions like first-year interns, a minimal level of pre-existing skills is necessary. Additionally, the skill sets required of trainees change rapidly in a short period of time. The skill set to be a good intern is not necessarily the same as a junior or senior resident, and high performers on an oncology rotation might have a much different set of knowledge, skills, and attitudes than a top performer on a trauma rotation. The ever-growing breadth and diversity of subspecialty training

available after residency programs is a key indicator that widely different interests and skills sets are needed in modern medical practice. Indeed, proponents of SBS might overlook the raison d'être of training programs—to impart the necessary competencies required for later practice. Focusing efforts on investigating applicant ability to perform duties that can be trained might be of little value. Programs might be wise to focus their efforts on developing in-training programs based on best practices, such as deliberate practice,[26] competency-based progression,[27] and adequate feedback and debriefing.[28] If we want high performers in medical training programs, the answer is not to seek out those who already have those skills, the solution is to develop high-quality and comprehensive training programs that can develop a diverse group of learners. However, it might be more logical for programs to assess abilities and competencies that are less malleable to training and unlikely to be remediated, such as integrity, personality, etc.

Resources

Obviously, one of the biggest concerns about implementing SBS systems is the time, money, and resources that are required to develop, implement, and assess such programs. Executing this selection strategy in the most appropriate and valid way would require a program to conduct a thoughtful job analysis ascertaining key competencies of the position and program, identify assessments to evaluate how to measure those competencies, develop testing procedures, select and/or train qualified assessors, and evaluate the success of those assessments.[29] This process would require extensive use of both faculty and incumbent trainees to serve as subject matter experts and help validate potential simulation-based assessment tools. Additionally, many of the test development and assessment procedures would require institutions to have trained selection scientists on the payroll and/or outsource these jobs to external consulting companies. In addition, when an ideal solution is developed, programs must then find resources to implement the simulations among hundreds of applicants. Others have noted the vast financial costs associated with typical current selection systems that primarily entail application review and in-person interviews (AKG, BJD, unpublished data, 2016). Adding more simulations to that process will likely be a costly endeavor and terribly burdensome, especially if the solution is to have each applicant engage in technical skills demonstration, role play, team-based exercises, etc. Implementing online or paper-based situational judgment tests might be least costly, but still require the test development methodology noted here. In short, medical training programs do not have the funds available that other industries have to invest in SBS. Additionally, until additional research is conducted examining the validity of such practices in a medical education environment, programs might have little incentive to invest such funds if and when they were available.

Opportunities

Opportunities abound for additional research in this aspect of medical simulation. From the development and validation of baseline SJT paper tests to complex interactive team-based simulations, this aspect of medical simulation remains largely uninvestigated. The next steps require investigators to "start small" and develop generalizable resource-sparing evaluations and acquire psychometric evidence of validity. As programs begin to cater these simulations to their needs, they will inevitably develop more complexity and specificity to the specific competencies and subspecialties for which they were designed. Additionally, investigating the qualities and traits most likely to result in success in a predefined environment (eg clinical medicine) is also fruitful ground for additional investigation.

CONCLUSIONS

General surgery programs are receiving unprecedented numbers of applications for categorical residency positions. The goal of the selection process then is to identify the most competent candidates and those who are most likely to succeed will succeed in a specific program. By many current metrics (attrition, remediation, etc), though, our current selection process is not achieving this aim in the most efficient manner. Simulation offers many opportunities to serve as a powerful complement to this process because it allows program leadership to evaluate an applicant's performance on exercises that replicate as closely as possible different responsibilities and demands of the job. Additionally, participating in these exercises can offer applicants a realistic preview of what it would be like to be a trainee in the program. Future work is needed to explore which types of simulation activities are best for the desired competencies, examine the efficiency and practicality of incorporating these into the application process, and investigate the value added of each pathway via data analytics.

Author Contributions

Study conception and design: Gardner, Ritter, Paige, Ahmed, Fernandez, Dunkin

Acquisition of data: Gardner, Ritter, Paige, Ahmed, Fernandez, Dunkin

Analysis and interpretation of data: Gardner, Ritter, Paige, Ahmed, Fernandez, Dunkin
Drafting of manuscript: Gardner, Ritter, Paige, Ahmed, Fernandez, Dunkin
Critical revision: Gardner, Ritter, Paige, Ahmed, Fernandez, Dunkin

REFERENCES

1. Lievens F, Coestsier P. Situational tests in student selection: an examination of predictive validity, adverse impact, and construct validity. Int J Select Assess 2002;10:245–257.
2. Powis DA. Selecting medical students. Med Educ 1994;28:443–469.
3. Lievens F, Patterson F. The validity and incremental validity of knowledge tests, low-fidelity simulations, and high-fidelity simulations for predicting job performance in advanced-level high-stakes selection. J Appl Psychol 2011;96:927–940.
4. Schmitt N, Ostroff C. Operationalizing the "behavioral consistency" approach: selecting test development based on a content-oriented approach. Pers Psychol 1986;39:91–108.
5. US Merit Systems Protection Board. Job simulations: trying out for a federal job. Available at: http://www.mspb.gov/netsearch/viewdocs.aspx?docnumber=452039&version=453207&application=ACROBAT. Accessed February 5, 2016.
6. Boyce AS, Corbet CE, Adler S. Simulations in the selection context. In: Fetzer M, Tuzinski K, eds. Simulation for Personnel Selection. New York: Springer: New York; 2013:17–41.
7. Yaghoubian A, Galante J, Kaji A, et al. General surgery resident remediation and attrition. Arch Surg 2012;147:829–833.
8. Sanfey H, Williams R, Dunnington G. Recognizing residents with a deficiency in operative performance as a step closer to effective remediation. J Am Coll Surg 2013;216:114–122.
9. Makdisi G, Takeuchi T, Rodriguez J, et al. How we select our residents—a survey of selection criteria in general surgery residents. J Surg Educ 2011;68:67–72.
10. Longo WE. Attrition: our biggest continuing challenge. Am J Surg 2007;154:567–575.
11. Lewis MM. Moneyball: The Art of Winning an Unfair Game. New York: WW Norton; 2003.
12. Baumer B, Zimbalist A. The Sabermetric Revolution: Assessing the Growth of Analytics in Baseball. Philadelphia, PA: Penn Press; 2014.
13. Schmidt FL, Hunter JE. The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 85 years of research findings. Psychol Bull 1998;124:262–274.
14. Callinan M, Robertson IT. Work sample testing. Int J Select Assess 2000;8:248–260.
15. Earnest DR, Allen DG, Landis RS. Mechanisms linking realistic job previews with turnover: a meta-analytic path analysis. Pers Psychol 2011;64:865–897.
16. Hausknecht JP, Day DV, Thomas SC. Applicant reactions to selection procedures: an updated model and meta-analysis. Pers Psychol 2004;57:639–693.
17. Schaverien MV. Selection for surgical training: an evidence-based review. J Surg Educ 2016;73:721–729.
18. Fortune JB. The content and value of letters of recommendation in the resident candidate evaluative process. Curr Surg 2002;59:79–83.
19. Greenberg AG, Doyle J, McClure DK. Letters of recommendation for surgical residencies: what they say and what they mean. J Surg Res 1994;56:192–198.
20. Kim RH, Gilbert T, Suh S, et al. General surgery residency interviews: are we following best practices? Am J Surg 2016;211:476–481.
21. Schatz A. Football's Hilbert problems. J Quant Anal Sports 2005;1:2.
22. Wolfson J, Addona V, Schmicker RH. The quarterback prediction problem: forecasting the performance of college quarterbacks selected in the NFL draft. J Quant Anal Sports 2011;7:1–19.
23. Kuzmits FE, Adams AJ. The NFL Combine: does it predict performance in the National Football League? J Strength Cond Res 2008;22:1721–1727.
24. Yeo H, Bucholz E, Sosa JA, et al. A national study of attrition in general surgery training. Which residents leave and where do they go? Ann Surg 2010;252:529–536.
25. 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 May 17, 2016.
26. Ericsson A, Krampe RT, Tesch-Romer C. The role of deliberate practice in the acquisition of expert performance. Psychol Rev 1993;100:363–406.
27. Yudkowsky R, Park YS, Lineberry M, et al. Setting mastery learning standards. Acad Med 2015;90:1495–1500.
28. Rudolph JW, Simon R, Dufresne RL, Raemer DB. There's no such thing as "nonjudgmental" debriefing: a theory and method for debriefing with good judgment. Simul Healthc 2006;1:49–55.
29. Gatewood RD, Feild HS, Barrick M. Human Resource Selection. Mason, OH: South-Western; 2011.