Chapter 1
Simulations for Personnel Selection: An Introduction

Kathy Tuzinski

Multimedia simulations stand alone in their ability to elicit, capture, and measure the behaviors that are most similar to actual job performance. This book surveys the current landscape of multimedia simulations for personnel selection, with accessible chapters written by those who have shaped the landscape through their pioneering efforts to merge new technologies into the practice of industrial and organizational (I-O) psychology. It will be of great interest to students, researchers, and practitioners who are looking for guidance in developing and implementing multimedia simulations for employee selection. It is also a valuable source of information about the wide range of simulations in use today, and is designed to provide inspiration, ideas, and lessons learned to novice and expert simulation developers alike.

One can find books that address various concepts related to simulations, such as computer automated scoring and the use of technology in assessment, but this is the first book to focus entirely on multimedia simulations. Other books about simulations are geared towards the implementation, development and scoring of credentialing and licensing simulations (exams) that primarily measure field-specific knowledge and highly technical skills, but not necessarily the constructs that are more common in the workplace, such as personality, judgment, and other job-related 'soft' skills. This book treats the world of multimedia simulations for personnel selection as a discipline in its own right, incorporating wide-ranging issues such as implementation, scoring, development, and validation. With helpful chapters containing the best technology for developing simulations, step-by-step instructions and lessons learned, as well as latest research on candidate reactions and group differences, this book is likely to be the best resource available.

Today's multimedia simulations take a variety of forms. Easy to recognize, but difficult to define exactly, these multimedia simulations are numerous and challenging to categorize. The one common thread that runs through all simulations is their design for one purpose: to capture work-relevant performance, either while performing a task, interacting with another person, or working with systems. They also share an ability to keep candidates more engaged relative to other types of assessments.

K. Tuzinski (✉)
CEB, Minneapolis, MN, USA
e-mail: kathy.tuzinski@shl.com

From that common thread, however, simulations diverge. Organizing frameworks are lacking. An organizing framework would help not only to bring greater definition to the term "simulation" but also, may serve as a basis for prescriptive simulation design, laying the foundation of standards for simulation development and processes for validity, reliability, and scoring. It may also help the reader understand how the different examples in the book fit within the larger simulation space. Before offering a view of what an organizing framework might look like, it is helpful to make a brief visit to the roots of today's modern simulations.

1.1 Multimedia Simulations in Context

> Hopefully, these first few attempts are the beginning of a whole new technology of behavior sampling and measurement, in both real and simulated situations. If this technology can be realized and the consistencies of various relevant behavior dimensions mapped out, the selection literature can cease being apologetic and the prediction of performance will have begun to be understood – Wernimont and Campbell (1968, p. 376).

Written over 40 years ago, the above quote is prescient. Simulations have come a long way since their formal introduction into the military during the 1940s, and a decade later, when they figured prominently in the managerial selection and development programs at AT&T.[1] As technological innovations advanced, simulation developers found ways to capitalize on the opportunities. Today's simulations incorporate all forms of multimedia, including audio, video, and 3D animation, and automation in the delivery and scoring. Simulations are now used everywhere and for a wide variety of positions.

Simulations are best characterized as measurement methods, rather than a type of test or construct. In its various forms, simulations measure hard and soft skills, personality, task performance, job knowledge, and cognitive ability. Furthermore, simulations are used in a variety of applications, including certification testing (e.g., ophthalmic technicians), licensure, training, and personnel selection. Simulations are rooted in three categories of tools, which by themselves show similarities and overlap: (1) assessment centers, (2) work samples, and (3) situational judgment tests. At the risk of oversimplifying, the multimedia simulations contained within this book reflect these tools with technology added for the purposes of improving and automating assessment delivery, data capture, and scoring.

1.1.1 Assessment Centers

The assessment center method grew out of the research labs of the modern twentieth century psychological measurement movement and the field of military selection.

---

[1] Some good resources on the history of testing: reaching back to China 3000 years ago (Oakland 2006); the progression of testing from the birth of psychological science in the late nineteenth century up through modern testing (Waldron and Joines 1994); assessment centers (Lievens and Schollart 2009); situational judgment tests (Whetzel and McDaniel 2009); and early multimedia assessments (McHenry and Schmitt 1994).

Psychologists developed precursors to the modern assessment center method in their research (e.g., Henry Murray's Harvard Psychological Clinic Study in the 1930s) and applied testing in military contexts (German Officer Selection in 1920s to 1942; British War Office Selection Boards, and the United States Office of Strategic Services (US OSS) Program in World War II). Eventually these methods made their way into managerial selection and development in the business world. These methods were first applied to the civilian sphere in the British Civil Service Assessment in 1945 and AT&T Management Progress Study which started a decade later. The typical assessment center incorporates some or all of the following components: multiple measures, observations, and assessors; behavior capture; trained assessors; and the integration of behaviors by pooling information (ratings or impressions) from raters.

Assessment center exercises are simulations of major aspects of performance for a given role. Assessment centers have traditionally been used for assessing managers, but have also been used for sales people, and roles for public safety, for both selection and development. Exercises such as the in-basket, leaderless group discussion, and the role play are all meant to test out a candidate in a real life situation to look for samples of performance. Although they can be used for measuring many different types of job related skills and abilities, they tend to measure competencies—a mixture of knowledge, skills, abilities and other characteristics (KSAOs)— such as interpersonal skills, communication skills, planning and organizing, and analytical skills (SIOP 2013).

More than 20 years ago, Waldron and Joines (1994) made predictions about things to come in assessment centers, such as multiple-choice in-baskets, the increased use of low fidelity simulations, and remote assessment. Furthermore, they predicted an increase in automation, including simulated email systems, data integration, and exercise scoring and reporting. Today's managerial assessment centers in fact do incorporate the latest in technology that is reflected in today's jobs, providing a more realistic twenty-first century "day in the life" experience relative to traditional brick and mortar assessment centers. Text messages and emails have been incorporated into simulations, allowing for information to arrive non-sequentially, just like the way it is in the work day of the typical manager of today (McNelly et al. 2011). Demand for an alternative to the in-person assessment center has also increased in response to budgetary constraints and the increasing affordability of remote assessment and technology-enabled assessment. For those interested in getting a sneak peek into the future of assessment centers, the chapter by Guidry et al. on novel techniques for tracing decision-making process in managers will provide ample food for thought.

1.1.2 Situational Judgment Tests

Situational judgment tests (SJTs) appeared on the scene at about the same time as assessment centers. In the 1920s, a widely used SJT with response options was likely a subtest from the George Washington Social Intelligence Test (Whetzel and McDaniel 2009). Army psychologists used the SJT format for measuring the judgment of

soldiers in World War II, and a number of SJTs were developed in the 1940s for assessing supervisory potential. By the 1950s and 1960s, organizations were using SJTs for managerial selection (Whetzel and McDaniel 2009).

SJTs present situations to candidates that might be encountered on the job and ask candidates to respond in one of two ways, what they would do or what they should do, given the situation. Situational judgment tests have been used for measuring a number of different constructs, such as interpersonal (Lievens and Coetsier 2002) and leadership judgment (Bergman et al. 2006), and conflict resolution skills (Olson-Buchanan et al. 1998).

SJTS are preferred over higher fidelity simulations, such as assessment center role plays, because of their ease in administration and scoring. SJTs have grown in popularity for a number of reasons. First, their high face validity provides more favorable candidate reactions, relative to standard personality or knowledge-based multiple-choice or Likert-scaled formats. Second, smaller subgroup differences (i.e., minority-white, or female-male) have been found for some SJTs relative to traditional cognitive ability tests (Clevenger et al. 2001), thus providing an opportunity for greater validity while minimizing the risk of adverse impact. Third, large-scale studies indicate that SJTs have substantial criterion-related validities (McDaniel et al. 2001).

The contact center simulation featured in the Holland and Lambert chapter in this volume uses an SJT item format to gather candidate responses on the most effective response to say, given the situation, where the "situation" can be any combination of what the caller just said, the information provided in the simulated agent software, and other pertinent information that was provided to the candidate during the assessment. The coaching simulation described by Gutierrez and Meyer in this volume is an example of an SJT that uses a video-based format. In moving from a low-tech, in-person assessment center exercise (role play) to a high-tech remote situational judgment test (coaching simulation), the result is a lower-fidelity assessment that is lower in cost for organizations, has higher availability, standardization, and ease of use.

1.1.3 Work Samples

Definitions for work samples have varied in the literature and there appears to be some disagreement over what constitutes a work sample. Simply stated, this disagreement is over whether work samples include a wide range of measures at the high and low ends of the (primarily) physical fidelity spectrum, or whether this label is reserved only for those measures at the high end of the spectrum. The broader definition would include low-fidelity measures such as "talk through" interviews and situational judgment tests in addition to high-fidelity measures, such as cockpit simulators, performance tests, and assessment center exercises.

The most literal definition of a work sample would be a hands-on performance test in which a job applicant is required to actually perform a job-related task under the same conditions as those required on the job. Measures that are classed under the heading of work samples can be organized according to the degree by which they

are removed from the two features of actual hands-on performance and a real work setting (Callinan and Robertson 2000).

In the broader definition, there is room for work samples to include both situational judgment tests as well as assessment center exercises if level of fidelity is specifically called out in the definition of a work sample (as in "the contact center simulation is a low-fidelity work sample"). The narrower definition of work samples, on the other hand, would only include measures in which the applicant performs a selected set of actual tasks that are physically and/or psychologically similar to those performed on the job (Roth et al. 2005). This difference in opinion has created some confusion in the literature in terms of when a test is a work sample vs. something else, and how to interpret previous research findings, particularly surrounding the validity of work sample tests.

1.2 The Fidelity Continuum: An Organizing Framework

Simulations vary in their ability to replicate the physical and psychological fidelity of a work task. Physical fidelity is the extent to which a test itself involves the actual tasks performed on the job (Truxillo et al. 2004), whereas psychological fidelity represents the extent to which the relevant knowledge, skills and abilities (KSAs) are called forward in the process of completing the task (Goldstein et al. 1993). In both cases, the degree of fidelity is a continuum rather than a dichotomy.

Physical and psychological fidelity are related concepts: although some psychological fidelity comes along whenever there is physical fidelity, it is possible to have psychological fidelity without physical fidelity (think, for example, of a paper-and-pencil SJT). However, Goldstein et al. (1993) held that physical fidelity is less important to content validity than psychological fidelity, particularly if the job requirements do not involve physical-based tasks, such as operating machinery, fixing equipment, manipulating physical objects, and so on.

Fidelity maximizes the point-to-point correspondence between the simulation and the task it is meant to represent (Asher and Sciarrino 1974) and should therefore have a direct impact on the validity of the measure. It increases face validity, which can improve candidate's perceptions of the assessment as well as provide the benefit of a realistic job preview. Poor face validity has been suggested to reduce the candidate's desire to perform well, possibly leading to biased test scores (Arvey et al. 1990). Fidelity also aids content validity by matching the KSAs brought out by the simulation to the requirements of the role, which is important if this is the main strategy for justifying the use of the assessment (i.e., a selection procedure can be validated by a content-oriented strategy if it is representative of the important aspects of performance on the job, according to the Uniform Guidelines).

According to Lievens and De Soete (2012), the logic of maximizing point-to-point correspondence between the predictor and criterion is conceptualized differently for high-fidelity versus low-fidelity simulations. In high-fidelity simulations (assessment centers and work samples), assessors observe and rate actual on-going candidate

behavior, which shows true point-to-point correspondence with the criterion. On the other hand, low fidelity simulations (such as an SJT) sample applicants' procedural knowledge about effective and ineffective courses of action in job-related situations. There is not the same level of point-to-point correspondence in low-fidelity simulations because the behavior of choosing among alternatives is not the same as constructing and actually demonstrating the behavior one wishes to make in response to complex interactions with other humans (Thornton and Rupp 2006). However, it is important to note that moving to low-fidelity forms does not necessarily in all cases harm the criterion-related validity (Lievens and Patterson 2011).

Many researchers have suggested that stimulus and response fidelity should be considered separately, particularly when interpreting research findings (Lievens and De Soete 2012; Truxillo et al. 2004). For example, Funke and Schuler (1998) found that moving towards higher fidelity on the stimulus side (from orally-presented questions to video) had little effect on the validity but moving to higher fidelity on the response side (multiple-choice vs. written vs. orally-given replies) did affect validity. Response fidelity appeared to put a ceiling on the gains in validity that could be achieved by increasing the fidelity of the stimulus. Given the proliferation of simulation types, with varying levels of fidelity on the stimulus and the response sides, treating stimulus and response fidelity separately is good advice. What started out as a suggestion, is now likely critical when interpreting research findings. It is expected that more research in this area will continue, with findings such as those recently found by Lievens et al. (2012), demonstrating that response fidelity may not only affect validity but also have modest effects on reducing the saturation of cognitive ability, increasing the saturation of certain personality traits, as well as improving candidate perceptions.

To provide an organizing framework around work samples, performance tests and competency testing, Truxillo et al. (2004) grouped these assessments into three main categories based on level of physical fidelity. The first group of tests included those that were physically just like the job. The second group of tests represented those that closely sampled the tasks performed on the job, such as the physical ability tests often used for selecting into public safely positions. The third group of tests were those that closely resembled the job (in that they present applicants with a work-related situation through video), but rather than applicants showing what they would do, applicants described what they would do in a given situation. This framework has been adopted here for grouping simulations into different fidelity categories, treating the stimulus and responses separately. Tables 1.1 and 1.2 represent a framework for simulations, based on where the stimuli (Table 1.1) and responses (Table 1.2) fall on a continuum of fidelity. For the purposes of simplicity, physical fidelity and psychological fidelity are not shown, but it may be assumed that psychological fidelity exists across all forms of stimuli, increasing from left to right, and physical fidelity is primarily present in the last column only.

This framework is useful as a starting point for organizing research findings on validity, applicant reactions, and group differences, and may be useful for a few different applications: to categorize and understand the differences among the possible simulations described in this book and elsewhere; to inform R&D spending by

Table 1.1 Continuum of stimulus fidelity

	Lower			Higher
Presentation	Written or oral delivery	Video, audio or animation (linear)	Video, audio or animation (branching)	Live action
Description	Computer, paper-and-pencil, or orally delivered situation	Stimuli (e.g., action, dialogue, or computer system screens) that do not adapt in response to the candidate's actions	Stimuli that change in response to the candidate's actions	Stimuli that maximize point-to-point correspondence between the task and the simulation and change in response to the candidate's actions
Example assessment types	Situation-based interviews, written SJTs	Call center simulations, interpersonal simulations, static role plays and in-basket exercises, static software simulations	Serious games, dynamic (computer adaptive) simulations	Work samples, flight simulators, role plays, in-baskets, job tryouts

Table 1.2 Continuum of response fidelity

	Lower			Higher
Description	Closed-ended (linear)	Closed-ended (branching)	Open-ended declarative	Open-ended behavioral
Response types	Multiple choice questions	Multiple choice questions	Written or oral response	Demonstrated behaviors

highlighting areas of greatest return on investment; and to provide a means for the practitioner to evaluate different simulation options when making decisions about which to use. For example, if research demonstrates that increasing response fidelity greatly improves candidate reactions and validity and decreases group differences, this would provide the practitioner with the evidence needed to build the case within the organization to invest in response-gathering technology, such as webcams, that can be provided to candidates during the assessment process (Oostrom et al. 2011). The technologies described by Guidry et al. in Chap. 11 for collection and analysis of free-form behavioral responses would likely get increased attention as a result of this research, especially if it can be coupled with the new scoring technologies described by Sydell et al. in Chap. 5.

The investment in technologies to increase fidelity is more easily justified for jobs where mistakes could be costly, dangerous, or fatal (such as pilots, air traffic controllers, and surgeons). However, many organizations in the civilian sector do not hire for such mission-critical roles. Should the same level of stringency be applied in the case of a front-line manager or call center agent? For these roles, costs of implementing the simulation are considered along with the cost of making a bad hire. In Chap. 2, Boyce et al. address fidelity in greater depth and discuss the benefits (and drawbacks) of fidelity depending on the intended purpose.

Level of fidelity is just one factor that differentiates simulations. Simulations vary in the constructs measured (or tasks represented), comprehensiveness (degree to which the entire job performance domain is represented by the tasks that make up the measure), job role, purpose (i.e., credentialing, training, or selection), and difficulty. As research findings on different simulations accumulate, there will be greater opportunities to assemble these findings in a framework that highlights the effects that various simulation facets have on important outcomes such as applicant perceptions, group differences, and validity.

1.3 Preview of Chapters in the Book

Creating a new simulation from scratch is as much an artistic as a scientific endeavor. An apt metaphor is that of trying to write a symphony or putting together the pieces of a 3D puzzle. There are many layers to consider: content, look-and-feel, flow, scoring, and multimedia elements. The final product needs to be psychometrically sound as well as provide an experience to the candidate that feels authentic and coherent and also has the ability to draw out the relevant performance behaviors.

It can be challenging for even the most talented divergent thinker to get his or her mind around the process. Many of the authors in this book are pioneers in this new area of test development. They have had to rely on their own ingenuity and skills in measurement, psychometrics, and storytelling without the benefit of how-to manuals or best practices for simulation development. This is why this book is so needed at this point in time.

1.3.1 Section I: Simulations in the Selection Context— Broader View

Section I introduces some important topics when considering the use of simulations for employee selection. The chapters in this section address a wide variety of issues, including challenges and opportunities facing simulations in the selection context, current and emerging tools and technologies available to simulation developers, methods for scoring simulations, and current and future research directions for candidate reactions to simulations.

Kicking off Section I, Boyce et al. in Chap. 2 discuss the considerations, challenges, and opportunities of simulations in the employee selection context. They cover the organizational and individual issues relevant to the use of simulations during the attraction and recruitment stages and beyond. The authors explore the existing research on each issue and provide practical guidance, highlight areas in need of additional research, discuss the ideal conditions for simulation use, and share strategies for effective implementation within organizational selection contexts.

Bruk-Lee et al. in Chap. 3 review the research to date on candidate reactions to simulations and media rich assessments. They begin broadly, with a historical overview of candidate reactions to assessment and then focus specifically on candidate reactions to technologically advanced assessments through the lens of procedural and distributive justice. Furthermore, they propose that such research should distinguish between administration mediums and media types, as well as consider the impact of individual differences. Finally, the concept of the uncanny valley (candidates' perceiving animations as 'eerie' or 'unhumanlike') is explored. They highlight the paucity of research in candidate reactions to media rich assessments in general and call for an increased emphasis on this line of research.

Hawkes in Chap. 4 provides an overview and evaluation of the tools available to assessment simulation creators. Written for the non-technical reader, this chapter provides a how-to guide for selecting the most appropriate tools for simulation projects. It addresses each stage of the simulation production process, covering technologies for content creation, authoring, and deployment. Hawkes provides additional insight into the important things to consider when deciding on which technology to choose, including ease-of-use, flexibility, limitations, accessibility, and cost. Real-world examples of these technologies are sprinkled throughout the chapter, and a comprehensive list of links to additional resources is provided.

Closing out Section I, Sydell et al. in Chap. 5 review the scoring of simulations. They believe that harnessing the power of larger data sets and advancements in the

ability to combine items and item types into scoring algorithms will bring a shift in the predictive capabilities of simulations. They also discuss the use of automatic scoring and branching logic, as well as new technologies for automating the scoring of qualitative (i.e., open-ended) assessment responses. They assert that a much deeper understanding of who a person is and how they will behave can be developed by examining interactions between simulation components and other sources of information.

1.3.2   Section II: Simulations in Action

Section II provides real world examples of simulations in action. The chapters in this section provide examples and discuss simulations for service roles, manufacturing, contact centers, and managerial selection. Simulations for assessing computer proficiency and leadership and decision-making are also discussed.

Barr and Coughlin in Chap. 6 describe simulations that measure computer skills, from more generic computer proficiency to the specific (e.g., Microsoft Office). These simulations have the potential to be more ubiquitous as they are relevant across a wide variety of job roles/levels and industries. In fact, they are relevant for any job where computer skills are required. Unlike other simulations that only need to resemble the psychological fidelity of workplace systems (call center software comes to mind), software simulations need to represent the interface (physical fidelity) as close to the real thing as possible. The authors provide a helpful discussion surrounding the many decisions that have to be made when designing and scoring computer/software simulations.

Holland and Lambert in Chap. 7 focus on the use of multimedia simulations in contact centers. Contact centers are challenging work environments that require the employees to provide a high quality customer service or sales experience to customers while interacting with multiple computer programs, under significant time pressure, and often under the watchful eye of performance management systems that track their communication style, reliability, and performance. Contact center simulations not only provide a realistic job preview but create more a engaging candidate experience. They have evidence showing that contact center simulations may be the single best predictor of many job-specific metrics.

O'Connell et al. in Chap. 8 describe the use of interactive simulations in manufacturing settings. They mention several examples of interactive simulations that are currently in use by manufacturing organizations. Manufacturing is an interesting case in which 'simulation' can just as easily mean a multimedia-based assessment measuring targeted competencies, as a complex multi-workstation setting involving the actual equipment and processes encountered on the job. As manufacturing roles change to include more decision-making, multitasking, and collaboration, the use of multimedia simulations in these jobs is only expected to increase.

LaTorre and Bucklan in Chap. 9 review simulations for service roles. Service roles include positions in retail sales, customer service, and banking, where interpersonal

effectiveness is often a top job requirement. Focusing on the application of the assessment center methodology to high volume, non-managerial positions, they describe elements of a best-in-class assessment program, along with lessons learned and important pitfalls to avoid. The chapter also describes research approaches to establish the return on investment (ROI) of simulations, and techniques for communicating this value to the hiring organization.

Gutierrez and Meyer in Chap. 10 describe the use of multimedia simulations for selecting managers and front-line supervisors. Simulations for these positions have gained in popularity in recent years as organizations have learned that they can be a cost-effective alternative to in-person assessments. Two different simulations were developed simultaneously, a coaching skills role play and an inbox assessment. Careful attention was paid to representing the types of scenarios that managers encounter on a daily basis, such as coaching direct reports, prioritizing one's work and that of others, monitoring employees, and making decisions under pressure with limited information. They have found these simulations to be well-received by candidates and human resources (HR) recruiters alike.

Guidry et al. in Chap. 11 focus on simulations that reflect the realities of the postrecession economy, the unanticipated and complex situations that have become the exception rather than the norm for business leaders today. Their focus, as they say, is to 'stretch the boundaries of virtual simulations' functionality'. They propose novel ways to leverage technology to measure previously unobservable decision strategies to bring out into the open normally concealed thought processes of candidates. They predict that detecting and measuring what has been up to this point difficult to assess using more traditional technologies will inform both the science and the practice of simulations for assessing and developing leaders.

Fetzer in Chap. 12 takes a peek forward at the future of simulations for employee selection. He believes that much can be done to advance the technology to develop new simulations that will render simulations of tomorrow to look more like games from the entertainment industry than traditional tests of yesterday. "Serious games," developed primarily for the military and government, are picking up momentum in the civilian sector, and will provide greatly enhanced levels of user engagement, measurement opportunity (and complexity), test security, and positive business outcomes for the organizations who utilize them in their hiring processes. Fetzer predicts that the future of game-inspired assessments (GIAs) will take two distinct forms: those that are more like the casual games of today, and those that are as realistic as the then-current technology will allow. Either way, he has no doubt that these GIAs will become the standard for personnel selection in the not-too-distant future.

1.4 Concluding Thoughts

The chapters in this book reflect the latest thinking and research in multimedia simulations and should be of interest to experts as well as students and a general business audience, and represents an important step towards normalizing the presence of

simulations in the context of personnel assessment. We suspect that in only a few years what is being represented in this book as novel will become more commonplace, and may even appear outdated. The field continues to move forward as simulation developers push the envelope by incorporating new technologies, adopt methods from other fields outside of I-O psychology (like credentialing, licensure, and education), work with a diverse set of talented people from marketing, technology, and multimedia production, and improve on scoring processes. It is personally a very gratifying time to work in this field.

References

Arvey, R. D., Strickland, W., Drauden, G., & Martin, C. (1990). Motivational components of test taking. *Personnel Psychology, 43,* 695–716.

Asher, J. J., & Sciarrino, J. A. (1974). Realistic work sample tests: A review. *Personnel Psychology, 27,* 519–533.

Bergman, M. E., Drasgow, F., Donovan, M. A., Henning, J. B., & Juraska, S. E. (2006). Scoring situational judgment tests: Once you get the data, your troubles begin. *International Journal of Selection and Assessment, 14,* 223–235.

Callinan, M., & Robertson, I. T. (2000). Work sample testing. *International Journal of Selection and Assessment, 8,* 248–260.

Clevenger, J., Pereira, G. M., Wiechmann, D., Schmitt, N., & Harvey, V. S. (2001). Incremental validity of situational judgment tests. *Journal of Applied Psychology, 86,* 410–417.

Funke, U., & Schuler, H. (1998). Validity of stimulus and response components in a video test of social competence. *International Journal of Selection and Assessment, 6,* 115–123.

Goldstein, I. L., Zedeck, S., & Schneider, B. (1993). An exploration of the job analysis- content validity process. In N. Schmitt & W. Borman (Eds.), *Personnel selection in organizations* (pp. 2–34). San Francisco: Jossey-Bass.

Lievens, F., & Coetsier, P. (2002). Situational tests in student selection: An examination of predictive validity, adverse impact, and construct validity. *International Journal of Selection and Assessment, 10,* 245–257.

Lievens, F., & De Soete, B. (2012). Simulations. In N. Schmitt (Ed.), *Oxford handbook of personnel assessment and selection* (pp. 383–410). New York: Oxford University Press.

Lievens, F., & Patterson, F. (2011). 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. *Journal of Applied Psychology, 96,* 927–940.

Lievens, F., & Schollart, E. (2009). Assessment centers: Overview of practice and research. Online readings in testing and assessment. http://www.intestcom.org/Publications/ORTA/ The%20assessment%20center.php. Accessed 28 Feb 2013.

Lievens, F., De Corte, W., & Westerveld, L. (2012). Understanding the building blocks of selection procedures: Effects of response fidelity on performance and validity. *Journal of Management.* doi:10.1177/0149206312463941.

McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology, 86,* 730–740.

McHenry, J. J., & Schmitt, N. (1994). Multimedia testing. M. G. Rumsey, C. B. Walker, & J. H. Harris (Eds.), *Personnel selection and classification* (pp. 193–232). Hillsdale: Lawrence Erlbaum Associates, Inc.

McNelly, T., Ruggeberg, B. J., & Ray Hall, C. Jr. (2011). Web-based management simulations: Technology-enhanced assessment for executive-level selection and development. In N. T. Tippins & S. Adler (Eds.), *Technology-enhanced assessment of talent* (pp. 253–266). San Francisco: Jossey-Bass.

Oakland, T. (2006). Introduction: The International Test Commission and its role in advancing measurement practices and international guidelines. In D. Bartram & R. K. Hambleton (Eds.), *Computer based testing and the internet: Issues and advances* (pp. 1–11). Chichester: Wiley.

Olson-Buchanan, J. B., Drasgow, F., Moberg, P. J., Mead, A. D., Keenan, P. A., & Donovan, M. A. (1998). Interactive video assessment of conflict resolution skills. *Personnel Psychology, 51,* 1–24.

Oostrom, J. K., Born, M. P., Serlie, A. W., & van der Molen, H. T. (2011). A multimedia situational test with a constructed-response format: Its relationship with personality, cognitive ability, job experience, and academic performance. *Journal of Personnel Psychology, 10,* 78–88.

Roth, P. L., Bobko, P., & McFarland, L. A. (2005). A meta-analysis of work sample test validity: Updating and integrating some classic literature. *Personnel Psychology, 58,* 1009–1037.

SIOP (2013). Types of employment tests: Assessment centers. https://www.siop.org/workplace/employment%20testing/testtypes.aspx. Accessed 7 Mar 2013.

Thornton, G., & Rupp, D. E. (2006). *Assessment centers in human resource management: Strategies for prediction, diagnosis, and development.* Mahwah: Lawrence Erlbaum Associates.

Truxillo, D. M., Donahue, L. S., & Kuang, D. (2004). Work samples, performance tests, and competency testing. In M. Hersen (Ed.), *Comprehensive handbook of psychological assessment* (pp. 345–370). Hoboken: Wiley.

Waldron, B., & Joines, R. (1994). Introduction to assessment centers: Assessment centers 101. Workshop conducted at the IPMAAC conference on public personnel assessment, Charleston, SC. June 26, 1994. http://www.ipacweb.org/files/ac101.pdf. Accessed 7 Mar 2013.

Wernimont, P. F., & Campbell, J. P. (1968). Signs, samples, and criteria. *Journal of Applied Psychology, 52,* 372–276.

Whetzel, D. L., & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review, 19,* 188–202.