A Century of Selection

Ann Marie Ryan[1] and Robert E. Ployhart[2]

[1]Department of Psychology, Michigan State University, East Lansing, Michigan 48824; email: ryanan@msu.edu

[2]Darla Moore School of Business, University of South Carolina, Columbia, South Carolina 29208; email: ployhart@moore.sc.edu

Annu. Rev. Psychol. 2014. 65:693–717

First published online as a Review in Advance on September 11, 2013

The *Annual Review of Psychology* is online at http://psych.annualreviews.org

This article's doi: 10.1146/annurev-psych-010213-115134

Keywords

employee selection, validation, technology and selection, selection constructs, selection methods

Abstract

Over 100 years of psychological research on employee selection has yielded many advances, but the field continues to tackle controversies and challenging problems, revisit once-settled topics, and expand its borders. This review discusses recent advances in designing, implementing, and evaluating selection systems. Key trends such as expanding the criterion space, improving situational judgment tests, and tackling socially desirable responding are discussed. Particular attention is paid to the ways in which technology has substantially altered the selection research and practice landscape. Other areas where practice lacks a research base are noted, and directions for future research are discussed.

Contents

INTRODUCTION

Whereas the scientific study of employee selection is now a century old, the practice of selecting employees is much more ancient. One of the first books on employee selection was published in 1913[1]: Hugo Munsterberg's *Psychology and Industrial Efficiency*, which described the selection of streetcar motormen, ship officers, and telephone switchboard operators. Perhaps because of its age, the study of employee selection is easily thought of as a mature research area. A vast number of studies on the topic have been published in addition to books, handbooks, and reviews; many psychologists are engaged full time in thriving practices associated with designing, implementing, and evaluating selection practices; and laws and professional guidelines direct the appropriate conduct of employee selection practice. With a century of research and psychologically driven practice, one may wonder: What is left to uncover? Haven't all the "big" questions been resolved?

What is fascinating and motivating to us is that plenty is going on. The area was last reviewed in the *Annual Review of Psychology* six years ago (Sackett & Lievens 2008). Today, the field of selection research is

(*a*) full of controversies, such as:

- Is the validity of integrity testing overstated (Ones et al. 2012, Van Iddekinge et al. 2012)?
- Are content validation strategies useful (Binning & LeBreton 2009, Murphy 2009, Murphy et al. 2009, Thornton 2009)?

(*b*) actively engaged in revisiting many "settled" questions, such as:

- Does differential validity exist for cognitive ability tests (Aguinis et al. 2010, Berry et al. 2011, Meade & Tonidandel 2010)?

[1]The German language version was published in 1912 and the English language version in 1913.

- What are the magnitudes of subgroup differences on commonly used predictors (Bobko & Roth 2013, Dean et al. 2008, Foldes et al. 2008, Whetzel et al. 2008)?
- Is there value in considering specific abilities above *g* in predicting job performance (Lang et al. 2010, Mount et al. 2008)?
- Are vocational interests of any value in selecting employees (Van Iddekinge et al. 2011a)?

(*c*) still working on major "intractable" challenges, such as:

- Can we reduce adverse impact in cognitive ability testing without sacrificing validity (DeCorte et al. 2010, Kehoe 2008, Ployhart & Holtz 2008, Pyburn et al. 2008)?
- Can we design personality tools for selection use that are not fakeable, or can we otherwise prevent or detect faking (Fan et al. 2012, van Hooft et al. 2012)?
- How can we convince stakeholders of the need to move away from subjectivity in hiring (Highhouse 2008, Kuncel 2008, O'Brien 2008)?

(*d*) expanding into literatures and organizational levels far removed from those historically investigated, including:

- How do we link selection methods and constructs to indices of unit-level outcomes (performance, turnover) (Van Iddekinge et al. 2009)?
- How might psychological research connect to firm-level strategy, economics-derived research on human capital resources, or competitive advantage (Ployhart & Moliterno 2011)?

(*e*) constantly being pushed by those in practice, who continually are confronting questions to which researchers have not yet produced answers (or haven't even begun to study!).

- Is unproctored, Internet testing appropriate, or are there major problems with this practice (e.g., unstandardized testing environments, cheating) (Beaty et al. 2011, Tippins 2009 and ensuing commentaries)?
- How well do prototypical assessment tools generalize globally (Ryan & Tippins 2009)?
- How can we design psychometrically sound tools to leverage new technological advances (e.g., immersive simulations) (Scott & Lezotte 2012)?
- How can we best measure knowledge, skills, abilities, and other characteristics (KSAOs) that have taken on increased importance in a rapidly changing work environment (e.g., adaptability, cross-cultural competence) (Inceoglu & Bartram 2012)?

So, as a mature field, selection research is not in a rocking chair in a retirement home but is more akin to a highly active senior who has not been slowed down by age.

Since the last *Annual Review of Psychology* article on this topic, multiple comprehensive handbooks on selection have been published (e.g., Farr & Tippins 2010, Schmitt 2012), new text editions specifically on the topic have emerged (e.g., Guion 2011), volumes have summarized knowledge of specific major topics (e.g., adverse impact, Outtz 2010; technology and selection, Tippins & Adler 2011), and high-quality review articles have been produced (Breaugh 2009, Macan 2009, Thornton & Gibbons 2009). Thus, there are many new summaries of the state of research, and producing a summary of the summaries would not be a useful contribution. Instead, we set our goal for this review to be more forward looking than backward looking and to zero in on those specific questions and topics that are currently of great interest as well as those that are likely to occupy research attention in the decade ahead. To base these predictions on the scholarly literature is a fairly obvious task. However, given the applied nature of selection research, we undertook a short survey of selection "practice leaders" (i.e., a dozen organizational psychologists whose full-time focus is on employee selection and who are viewed as influential in this arena). We mention views of these practice leaders throughout in discussing what is missing in the literature.

To provide some structure to our review, we have organized our discussion of research into the three areas of designing, implementing, and evaluating selection systems. To be sure, research covered in one section often has implications for the others, but some organizing structure is required.

DESIGNING SELECTION SYSTEMS

Selection researchers have long focused on two key questions in designing selection processes: What should be assessed? How should we assess it? In this section we review advances in considering what should be assessed (definitions and measures of desired outcomes of selection as well as constructs of focus in selection) and how it is assessed (methods of assessment). Although decisions regarding what to assess are based on job analysis, competency modeling, and other types of needs assessment, those topics have recently been reviewed in another *Annual Review of Psychology* article (Sanchez & Levine 2012) and thus are not discussed here.

Outcomes of Selection

For a number of years, selection researchers have noted the importance of "expanding the criterion space" (see Campbell 1990) or the need to define success at work more broadly than just task performance. In recent years, we have seen continued discussion of the need for expansion but also some action in terms of studying criteria other than task performance. There are several themes to this research.

First, more studies have looked at broadening the criterion space at the individual level. For example, predicting turnover with selection tools has been a greater focus (e.g., Barrick & Zimmerman 2009, Maltarich et al. 2010), and meta-analyses have examined relations between personality and organizational citizenship behavior and counterproductive work behavior (Chiaburu et al. 2011, Le et al. 2011, O'Boyle et al. 2012). However, given the focus on adaptive performance in today's workplace (Dorsey et al. 2010), it is surprising how little selection-related work has occurred. Therefore, although we see some expansion in what criteria are being studied by selection researchers, a continued expansion in the decade ahead is needed to improve predictive accuracy.

Second, selection researchers have begun to predict performance at unit and organizational levels rather than just at the individual level. **Figure 1** illustrates how multilevel considerations factor into selection research, with an indication of the amount of research in various areas. Units that hire more employees using valid selection scores show more effective training, customer service performance, and financial performance over time than units that do not (Ployhart et al. 2011, Van Iddekinge et al. 2009). Ployhart and colleagues (2009) found that positive changes in unit-level service-orientation KSAO scores generated corresponding positive changes in unit-level financial performance over time, but with diminishing returns. A meta-analysis by Crook and colleagues (2011) found that unit-level KSAOs are related to a number of unit-level performance metrics. This research is demonstrating that the tools and KSAOs used in selection may also relate to higher levels of financial performance, but the relationships may be more variable (contextualized) than found at the individual level. As the push for demonstrating the value of selection tools is likely to be as strong as ever in the years ahead, we anticipate work on the value of selection to higher-level outcomes to continue to emerge as an important area of research focus. Also, as the practice leaders we surveyed noted, our move away from a focus on the "job" as the unit of analysis and the dynamic nature of work role boundaries should push us to new and better ways of defining the desired outcomes of selection processes.

Third, we see some efforts to align selection systems with organizational strategy. One example of this would be cases where competency modeling processes are used to align selection

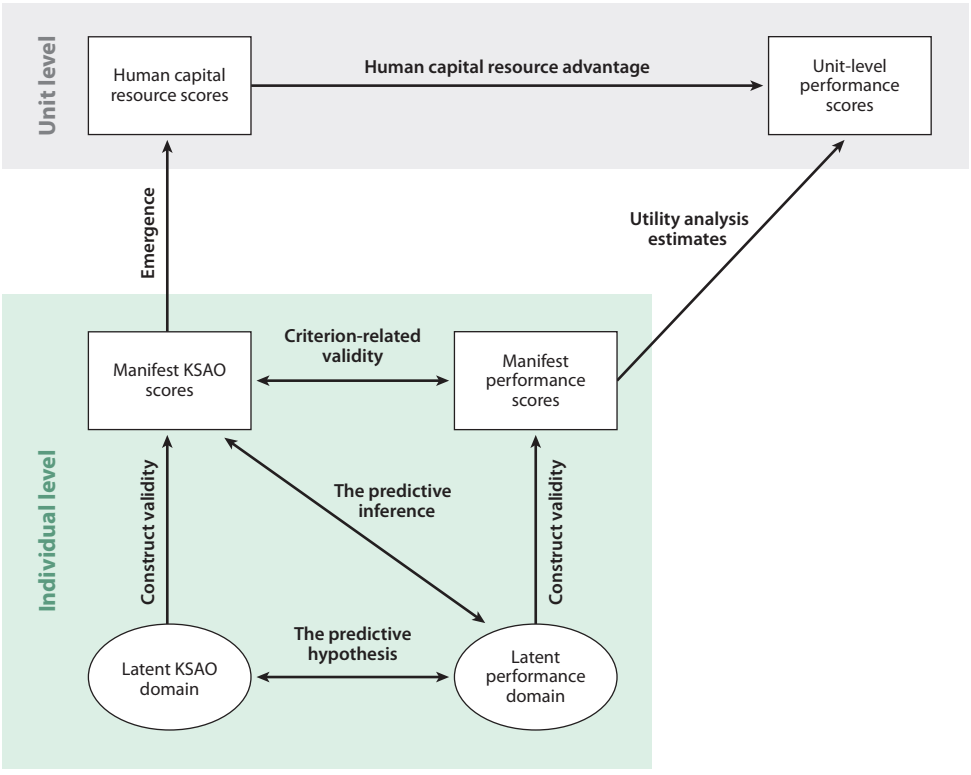

Figure 1

Illustration of multilevel relationships investigated in selection research. Abbreviation: KSAO, knowledge, skills, abilities, and other characteristics.

systems with other human resource (HR) systems and with the overall direction of the organization (Campion et al. 2011, Schippman 2010). Another example would be decisions regarding outsourcing of selection practices (Ordanini & Silvestri 2008). A final area links the KSAOs that are the target or focus of selection systems with an organization's human capital resources that underlie competitive advantage (Ployhart & Moliterno 2011, Wright & McMahan 2011). Human capital resources may generate firm-level competitive advantage because they are heterogeneous and difficult to imitate. Because KSAOs comprise the microfoundation of these human capital resources, selection may thus have a strong influence on a firm's ability to generate strategic competitive advantage (Coff & Kryscynski 2011).

What are we not doing? Cleveland & Collela (2010) argue that we are still limited in definitions of success and should include concepts such as employee health and well-being and nonwork criteria such as levels of work-family conflict. These constructs have been of interest to other psychologists, and even economists, for a long time, yet have not seemed to penetrate the criterion barrier in the selection field. At the same time, Cleveland & Collela (2010) note that the use of health status or family satisfaction as measures for selecting employees might create issues related to employee privacy and discrimination. We suggest a step back from a definition of success as including nonwork outcomes to a focus on work behaviors that lead to those outcomes. As an example, rather than attempting to develop selection tools to predict work-nonwork conflict, which could lead to assessment of "inappropriate" constructs (marital status, relationship stability),

selection designers might attempt to map out predictors of ability to set boundaries on work assignments, time management, and decision making in contexts of competing demands. Although these are not the only factors that contribute to work-nonwork conflict, they are ones that an employer might reasonably seek to predict and might be able to do so without straying outside what is acceptable.

As has been noted in selection reviews for most of the century of research, the "criterion problem" (Guion 2011) persists: Poor quality, contaminated, and mis- or underspecified measures of performance hinder our capacity to advance understanding of the true importance of individual differences as predictors. Although more data are now tracked by organizations (e.g., big data) on individual performance, we are still limited in our capacity to predict because of the challenge of obtaining accurate and complete assessments of individual behavior at work. However, with more different types of data available, it may be possible to assemble validity evidence based on multiple types of criteria, so that the limitations of one type are offset by the benefits of another type.

Constructs

Selection is essentially a predictive hypothesis, where one proposes that certain KSAOs of individuals are related to certain outcome(s) of interest to the organization (Guion 2011). This generally lends itself to looking at the prototypical individual differences (cognitive abilities, personality traits) studied in all areas of psychology as well as job-specific knowledge and skills. Rather than reviewing every study on specific KSAOs, we summarize the trends in what is assessed. In the next section we focus more specifically on the how, or the methods used to do that assessing.

Cognitive constructs. For many years, research on cognitive ability has been almost devoid in the selection area because it has been so firmly established as a predictor of job performance. As noted above, one ongoing development is that some researchers have questioned the blanket dismissal of the value of assessing specific abilities in conjunction with *g* (Lang et al. 2010, Mount et al. 2008) and are delineating contextual factors that might indicate when such a focus would be useful.

Scherbaum et al. (2012) recently argued that applied researchers need to give greater consideration to the developments in cognitive psychology and intelligence research more generally; others have suggested that the field is moving forward and is not missing much (Oswald & Hough 2012, Postlethwaite et al. 2012). Our own view is that there is room for expanding our focus from *g* to incorporating other cognitive psychology research into assessment. For example, there has been some interest in improving assessments of procedural knowledge, although most of this work is focused less on knowledge constructs and more on the methods used to assess this knowledge (see the section below on Situational Judgment Tests). Similarly, declarative knowledge assessment is not an area of research focus [for an exception, see Lievens & Patterson's (2011) examination of incremental validity of simulations over a declarative knowledge assessment]. The practice leaders we surveyed noted that they are often engaged in developing and implementing job knowledge assessments (e.g., programming, skilled trade knowledge), yet a lag exists in bringing findings from cognitive research on knowledge acquisition, situated cognition, and similar concepts (e.g., Breuker 2013, Compton 2013) into selection research. As another example, cognitive psychology work on multitasking and related concepts has yet to be broadly applied in selection contexts (for exceptions, see Colom et al. 2010 and Kantrowitz et al. 2012).

Noncognitive constructs. Active research is ongoing in the long-studied areas of personality and interest measurement, with a recent focus on the assessment of emotional intelligence, social skills, and integrity.

Personality. Over time, personality research in selection contexts has evolved from examinations of a hodgepodge of traits to a clear focus on the five-factor model constructs and, more recently, to a focus on using facets of the five-factor model for selection and/or looking at compound traits (e.g., proactivity, customer orientation) (for a review, see Hough & Dilchert 2010). Although recent studies have examined specific traits (e.g., affective orientation; Sojka & Deeter-Schmelz 2008) and nonlinear relationships between traits and outcomes (Le et al. 2011), selection-related personality research of late has a greater focus on how to assess than on what to assess, as we discuss below in the section on Self-Report Measures.

Interests. The value of vocational interests for use in employee selection has been reconsidered, and researchers have moved from prior views that interests are useful for occupational selection but that their range would be restricted and their value limited when selecting for specific jobs in specific organizations. Recent research suggests the usefulness of interests above and beyond personality measures (e.g., Van Iddekinge et al. 2011a) and for predicting both performance and turnover (Van Iddekinge et al. 2011b). We anticipate that this renewed attention to interests will result in valuable research in the next several years.

Emotional intelligence. A specific construct of focus in recent years has been emotional intelligence. Researchers have worked to better define and conceptualize emotional intelligence (Cherniss 2010, Harms & Crede 2010, Riggio 2010), to establish the predictive value of emotional intelligence (Blickle et al. 2009, Christiansen et al. 2010), and to document the relative roles of emotion perception, emotion understanding, and emotion regulation in predicting job performance (Joseph & Newman 2010, Newman et al. 2010). We anticipate that continued research will focus on social and emotional skills measurement in selection contexts for two reasons: Technology now allows for innovative and efficient ways of assessing these constructs, and much like personality inventories, these constructs are susceptible to socially desirable responding, making the design of useful tools a challenge (Lievens et al. 2011).

Integrity. Integrity test validity has been a focus of debate. Van Iddekinge and colleagues (2012) produced an updated meta-analysis of integrity test score criterion-related validities, finding them to be considerably smaller than prior research suggested. The findings from this study were in turn challenged for not being sufficiently comprehensive (Ones et al. 2012). At present it appears that integrity test score validities are large enough that they can be useful for selection, but the specific magnitude of these validities depends on which assumptions one is willing to make and which data sources one is willing to accept (Sackett & Schmitt 2012).

This debate has cast attention on definitions of rigor in primary study validation efforts as well as in meta-analyses and on how the goals of science and the commercial uses of science do not always align (Banks et al. 2012, Kepes et al. 2012). These are issues with broader implications for all of applied science, not just selection researchers. Although our practice survey participants did not single out integrity assessment as an area of research need, the focus in the business world on ethics (e.g., Drover et al. 2012) translates into a continued need to improve methods of integrity assessment. Much like with personality assessment, key challenges remain related to how to do so, given the issue of socially desirable responding in selection contexts.

Other constructs. One final note as to constructs of interest is the practice demand for tools that assess culture fit or person-organization fit. Although there is active research on person-organization fit (e.g., Meyer et al. 2010) and organizations adopting fit assessment tools, we

have not seen published work on the use of such fit assessments in selection contexts. This may reflect that many of these assessments are uniquely designed for particular organizations and their cultures, or it may indicate that they are simply personality and value instruments with a different label.

Methods

Arthur & Villado (2008) make the point that selection researchers compare apples and oranges by comparing research on a construct (e.g., conscientiousness, cognitive ability) with research on a method [e.g., interview, situational judgment test (SJT), biodata]. One emerging trend is greater attention to multiconstruct methods, particularly interviews, assessment centers, and SJTs; we discuss each of these in turn. Although little novel research specifically on biodata has been undertaken (however, see interesting work on elaboration by Levashina et al. 2012 and on scoring by Cucina et al. 2012), researchers continue to investigate ways to improve the use of personality assessment, which has applicability to all self-report inventories, including biodata measures. We discuss the issue of socially desirable responding and self-report measures as well.

Interviews. As the most commonly used selection tool, interviews are likely to be a target of research. At the time of the Sackett & Lievens (2008) review, there was an active research focus on interview structure and on what constructs interviews measure. Within the past six years, there actually has not been much research activity beyond some studies on acceptance and improving the value of structured interviews (Chen et al. 2008, Klehe et al. 2008, Melchers et al. 2011) and some focus on what constructs interviews measure (Huffcutt 2011). Research has instead shifted to impression management in the interview, with researchers emphasizing that impression management is not just a bias but is a contributor to the criterion-related validity of the interview (Barrick et al. 2009, 2010; Kleinmann & Klehe 2011; Swider et al. 2011). Research has also focused on the factors that influence initial impressions in the interview and how these initial impressions affect interview scores (Stewart et al. 2008).

Where should interview research head? In a review paper, Macan (2009) argued that we need to understand what constructs are measured best via the interview versus other methods, and we echo that as an important research need. Although our practice survey did not indicate demands from the field for interview research, given the increased use of technology in interviewing to reduce costs (Skype, videoconferencing; Baker & Demps 2009), research on these forms of interviewing is needed. Our personal experiences suggest that, much like with Internet testing, giving applicants responsibility for their own assessment environment does not mean they will make wise choices (e.g., how close to get to the camera, what angle to sit at, what is visible in the background) or have good environmental control (e.g., noise, interruptions), and these factors can negatively affect evaluations. In our experience, many firms have used the interview to assess fit with the company's culture and values. However, the fit with values rarely factors into interview selection research, so there is a missed opportunity to more directly influence interview practice.

Assessment Centers. Assessment centers are predictor methods that present the applicant with a variety of activities or exercises (e.g., inbox, role play, leaderless group projects) designed to assess multiple KSAOs. Despite several decades of activity, research on assessment center construct validity continues to sort through questions of what assessment centers actually measure and how to enhance their value (Dilchert & Ones 2009, Gibbons & Rupp 2009, Hoffman & Meade 2012, Hoffman et al. 2011, Jones & Born 2008, Meriac et al. 2009, Putka & Hoffman 2013; see also a 2008 series of articles in the journal *Industrial and Organizational Psychology: Perspectives on Science*

and Practice). One interesting result of this scrutiny is a better recognition of the ability to assess situational demands as a potentially important predictor of performance on a number of selection tools (e.g., assessment centers, interviews, SJTs; Jansen et al. 2012). Other suggestions emerging for practice are to use role-player prompts (Schollaert & Lievens 2012), reduce demands on assessors (Melchers et al. 2010), construct parallel forms of exercises (Brummel et al. 2009), and use technology for assessment exercise delivery (Lievens et al. 2010). We conclude that assessment center research remains vibrant and that many of the products of this effort have applicability to other selection methods (e.g., one can have parallel versions of interviews or SJTs, the demands on evaluators in interviews have similar influences as those on assessors, and construct validity questions similarly occur for SJTs and interviews).

Situational Judgment Tests. SJTs are predictor methods that present candidates with work-related situations and then assess how they might respond to those situations. Research interest in SJTs has grown to the point where there is now a reasonably large literature and corresponding narrative reviews (e.g., Lievens et al. 2008, Whetzel & McDaniel 2009). The majority of research published on SJTs since 2008 has followed three major themes.

First, there continues to be interest in examining the factors that may influence the criterion-related validity of SJTs. Several studies have demonstrated that SJT scores manifest validity across different high-stakes contexts (Lievens & Patterson 2011, Lievens & Sackett 2012). Other studies have examined the impact of response instructions on validity with Lievens et al. (2009) and have found no meaningful differences between two response instructions (what one would do versus what one should do) for predicting medical school performance. McDaniel and colleagues (2011) found that score adjustments that affect scatter and elevation can actually enhance validity. Motowidlo & Beier (2010) presented an alternative scoring system and found that scores based on this approach produced validity based on implicit trait policies and the knowledge facet of KSAOs.

Second, research continues to explore the nature of constructs assessed by SJTs. The construct validity of SJT scores has posed something of a mystery: Although scores on SJTs predict job performance, they do not fall into homogeneous groupings that correspond to existing KSAOs. A meta-analysis by Christian and colleagues (2010) classified SJT constructs into four broad categories: knowledge and skills, applied social skills (including interpersonal, teamwork, and leadership skills), basic tendencies (including personality), and heterogeneous composites of multiple KSAOs. Their review suggests most SJTs are intended to measure interpersonally related KSAOs, primarily social skills, teamwork, and leadership. One-third of the studies they reviewed failed to report the nature of any constructs measured by the SJT. Importantly, they found that matching the constructs assessed by the SJT to the nature of performance in a given context can enhance criterion-related validity. Other research has extended SJTs to measure different construct domains, including personal initiative (Bledow & Frese 2009) and team role knowledge (Mumford et al. 2008). In a different examination of construct validity, MacKenzie and colleagues (2010) found that the construct validity of SJT scores may be affected by whether the scores are obtained in applicant versus incumbent contexts. Finally, Westring et al. (2009) examined the systematic variance attributable to KSAO and situational factors and found that situations explained three times as much SJT score variance as did KSAO factors.

The final area of research has examined the magnitude of subgroup differences on SJT scores. Whetzel and colleagues (2008) found that whites perform better than African Americans, Asians, and Hispanics, and women perform better than men. Not surprisingly, the racial subgroup differences are larger when the SJT is more cognitively loaded. Bobko & Roth (2013) reported larger white/African American subgroup differences for SJTs than was found in previous research;

the disparities were credited to their consideration of constructs, context (applicant versus incumbent), and statistical artifacts.

Overall, research on SJTs continues to expand into new directions. SJTs have clearly established themselves as viable predictors of outcomes in a number of contexts. We suspect the next wave of this research will extend these findings into new areas, including predicting performance in other outcome domains (e.g., citizenship behaviors, turnover), employing SJTs in new cultural contexts, using SJTs to assess different constructs, and examining the effects of faking or applicant perceptions on SJT scores. Research that examines ways of enhancing the validity, reducing subgroup differences, or increasing efficiency will continue to be popular. Technology and SJTs are on a collision course, as most of the digital hiring simulations are in many ways more elaborate SJTs. At the same time, there are areas that have probably seen enough research. At present these areas include the effects of response instructions on SJT scores, and criterion-related validity studies in commonly examined contexts.

Self-report measures and socially desirable responding. The persistent problem of socially desirable responding or faking in applicant settings has continued to capture research attention. Some voices continue to call for a shift in focus from "faking" to developing broader theories of self-presentation or motivation (Ellingson & McFarland 2011, Griffith & Peterson 2011, Marcus 2009, Tett & Simonet 2011) that do not assume that such behaviors by applicants are illegitimate or have no positive value for prediction. There continue to be studies that address whether faking is a problem (Komar et al. 2008) and that examine the detection and magnitude of faking effects (Berry & Sackett 2009, LaHuis & Copeland 2009), but researchers have moved on to tackle the more thorny issue of how to reduce or eliminate faking. Approaches to reduction include warnings during testing (Fan et al. 2012, Landers et al. 2011) and eye tracking (van Hooft & Born 2012). Some of the most interesting work has revolved around variations on forced choice methods (Converse et al. 2008, 2010); for example, the US Army has adopted a multidimensional item response theory approach to personality assessment to address faking concerns (Stark et al. 2012). We expect the search for ways to use self-report tools more effectively in selection contexts will continue, as these tools are often cost effective and efficient for screening and add incremental validity, despite problems with socially desirable responding. However, without a common theory or even a definition of faking, these approaches will likely be piecemeal and incomplete.

Other methods. One method that seldom gets much research focus is the use of individual psychological assessment. Although some commentary has been generated recently around the need for more research on this practice area (see Silzer & Jeanneret 2011 and accompanying commentaries), there is not much to report in the way of empirical research on individual assessment. The validity of credit scores was supported by Bernerth and colleagues (2012), and some work has been done on what inferences are made in resume screening (Derous et al. 2009, 2012; Tsai et al. 2011). With the rise of e-portfolios (collections of evidence documenting competencies) for use in selection in certain sectors (e.g., education; Ndoye et al. 2012) and video resumes (Eisner 2010), investigation of what is being assessed by these methods as well as how to bring standardization to these evaluations is needed. Talk of using neuroimaging, genetic testing, and other methods that represent advances in other areas of psychology remains just talk (Zickar & Lake 2011); their application to selection is likely to remain hypothetical because of concerns with acceptability, particularly with regard to privacy, ethicality, and legality.

From a practice perspective, there needs to be more comparative research on method variations. SJT research on scoring provides a good example of how "going into the weeds" on topics such as

instructions and scoring of response options may provide insight into what works best. A few years back similar scrutiny was given to how to ask interview questions (e.g., behavior description versus situational description; Klehe & Latham 2005) and what different elements of interview structure are essential (e.g., Chapman & Zweeg 2005), resulting in improved advice on best practices in interview design. From a practice perspective, differences in instructions, item wording, response formats, and the like are changes that can be implemented with relative ease, and their "value" is therefore easy to sell to organizational decision makers. Yet, we seldom investigate their effects on validity, adverse impact, efficiency, and applicant reactions—we can and should do more. However, for this research on method variations to be published, it will be vital for researchers to develop a solid theoretical basis for proposing any differences and then to follow through on programmatic research using designs that strongly test the theory.

Combining Assessments

One final design topic relates to how one puts it all together. That is, besides deciding what to measure and how to measure it, selection system designers decide how those measures will be combined and implemented (whether compensatory or noncompensatory use; single or multistage; and any sequencing, weighting, and cut score decisions; DeCorte et al. 2011). There has been considerable revisiting of prior thinking in this area as well. For example, Hattrup (2012) summarized how ways of forming predictor composites are shifting as our view of the criterion space is expanding. Finch and colleagues (2009) considered how predictor combinations in multistage selection affect expected performance and adverse impact. DeCorte and colleagues (2011) provided tools for using pareto-optimal methods to determine the best ways to combine and use predictors with considerations of validity and adverse impact as well as cost and time constraints.

In sum, recent research on combinations and decisions has shown that simple mechanistic ideas for how to address adverse impact (e.g., add a noncognitive predictor, change the order of assessment administration) have deepened appreciation for the complexity of selection system design. Further work on why decisions are enacted the way they are (e.g., stakeholder assumptions, timeliness of hiring concerns) can point to the directions most in need of research. We suspect, as has been true for the past decade, that efficiency plays a major role in choices that employers make, and any research that addresses ways to improve efficiency is of as much practical value as that focused on effectiveness.

IMPLEMENTING SELECTION SYSTEMS

Since the Sackett & Lievens (2008) review, the major topics in research on implementation have been the use of technology (which certainly is part of design as well), globalization, and retesting.

Technology

In several previous reviews, the use of technology in selection has received mention, but since 2008, there truly has been an explosion in how technology has changed and continues to change selection practice. Computerized assessments are now fairly mainstream for big organizations, and off-the-shelf computerized assessment tools are used by smaller organizations. International guidelines for computer-based and Internet-delivered testing are providing some direction for best practices (Bartram 2009). Applicant tracking systems are standard and affect the choices organizations make about what tools can be implemented. The use of technology in recruitment continues to expand (for a review, see Maurer & Cook 2011). However, our survey of practice leaders suggested that

a greater examination of the usefulness and validity of these high-tech and multimedia methods vis-à-vis more traditional formats is still strongly needed to better understand their utility. Potosky (2008) provided a conceptual framework for considering how media attributes (transparency, social bandwidth, interactivity, and surveillance) might affect the psychometric properties of measures; research using this or other theoretical frameworks might enable us to make more significant strides in understanding and best utilizing technological advances in selection contexts.

Perhaps the most controversial area related to technology and selection has been whether using unproctored Internet testing is appropriate (for a good overview of the issues, see Tippins 2009 and accompanying commentaries). Unproctored Internet testing and proctored tests seem to evidence similar validities (Beaty et al. 2011) and similar levels of cheating or response distortion (Arthur et al. 2010). Many organizations have advocated that computer adaptive testing be used when going unproctored, as this will allow for less item exposure and less likelihood of cheating. Others have advocated for the use of a proctored verification or score confirmation test for presumed passers whenever an initial unproctored test is used. However, surprisingly little has been published on these emerging practices (see Kantrowitz et al. 2011 on computer adaptive tests; Scott & Lezotte 2012 for a general review of web-based assessments). Our surveyed practice leaders clearly felt there was not enough research: They said they want to see more data on the validities of UITs and the pros and cons of verification testing as well as research on the next generation of technology delivery (e.g., mobile testing), and we have to agree. However, for this research to make a broad scientific contribution, it will be vital for investigators to demonstrate the construct validity of their assessments, use field experiments or quasi-experiments, and/or develop a solid theoretical explanation for why differences should (or should not) exist between the administration modes.

There is no question that social media is causing a revolution in recruiting. Social media refers generically to Internet-based platforms (e.g., LinkedIn, Facebook) that connect people and organizations synchronously and asynchronously. Social media offers a potentially powerful mechanism for sourcing active and passive candidates, and although it has been primarily discussed as a recruiting tool, it is also being used in hiring decisions. Brown & Vaughn (2011), Davison et al. (2011), and Kluemper & Rosen (2009) discussed a number of concerns with using social media for employee selection purposes. The major conclusion is that although social media platforms may offer some opportunities for enhancing the speed, efficiency, or effectiveness of selection tools and processes, a number of legal (e.g., discrimination, privacy) and scientific (e.g., validity) issues must be carefully considered (Zickar & Lake 2011). This is an area where practice far outpaces research, and to date there is no published research that examines the validity, subgroup differences, or stakeholder opinions of using social media for hiring decisions. We hope this will change dramatically by the time of the next review on selection.

The value of technology for assessment is mentioned in a number of areas, but little empirical work exists to help guide innovative use. For example, although automated item-generation and test-assembly methods have taken hold in the educational testing literature (Gutl et al. 2011, Luecht 2000), we do not see published work on their viability or use in selection, although we know some practice leaders are engaged in deploying such methods to reduce the challenges of maintaining uncompromised, high-quality item pools, particularly in the cognitive ability domain. We hear discussions of how useful computational approaches to content analysis could be, such as latent semantic analysis (examining the relative positions of units of text in a semantic space of *N* dimensions) and data mining (automated approaches to identifying patterns in data) (see Indulska et al. 2012 for a comparative presentation on these); however, we do not see any examples of their applicability in selection systems in the published literature. Virtual role-plays and other immersive simulations are being marketed for employee selection use, but the published literature offers little information on them (e.g., McNelly et al. 2011 on web-based management simulations,

Oostrom et al. 2012 on computerized in-baskets). Indeed, "gamification," or the use of computer simulation games for business-related purposes (including recruiting and selection), is emerging as a stable industry yet gives little to no attention to the scientific literature on psychometrics or assessment. So although we hear talk about how crowdsourcing and other advances from the use of algorithms in Internet searches and in financial circles will make their way into making predictions about job applicants, we have yet to see published work on their application in this context (for an example of crowdsourcing to develop assessments in educational contexts, see Zualkernan et al. 2012). These are the research frontiers at the moment, and they are areas that investigators must examine or else the field runs the risk of being marginalized by gaming and technology-focused developers.

Globalization

Our survey of practice leaders indicated that globalization of selection systems is a top concern in need of research, particularly with regard to cross-cultural equivalence of tools, cross-cultural acceptability of practices, and navigation of selection environments (e.g., legal and social norms) when taking a process global. Although a sizeable literature exists on selection for international assignments (for a review, see Caligiuri et al. 2009), there has been a surprising lack of published research on globalizing selection systems (for a review, see Carey et al. 2010). Reserves of talent do not exist equally around the globe, as birthrates are declining or slowing in many Western countries, whereas developing countries have increasing populations (see, e.g., PriceWaterhouseCoopers 2010). Economic and labor pressures in many developed countries have led their organizations to source talent in developing countries, where the labor is often cheaper. Although historically this sourcing was limited to nontechnical jobs, currently it is increasing for professional occupations as well (e.g., engineering, accounting). Yet only limited selection research considers cross-cultural issues (for a review, see Caligiuri & Paul 2010).

Specific papers appear here and there on the use of global versus local norms (e.g., Bartram 2008), on cross-cultural equivalence of specific measures (e.g., Bartram 2013), on validity generalization across cultures (e.g., Ones et al. 2010), and on applicant reactions across cultures (e.g., Ryan et al. 2009). Almost all of this work points positively to the cross-culture transportability of methods, with some tweaks and adjustments for cultural context, yet it seems those in practice feel that the research base is still lacking. Ryan & Tippins (2009) provide an overview and specific guidance for the design, conduct, and implementation of global selection systems. Their recommendations outline a number of directions for future research that need to be addressed if the scientific literature is to better guide practice.

One other point related to globalization came from our survey of practice leaders, who noted that language proficiency assessment seems to be outside the expertise of most selection researchers, and thus its implementation into selection systems, validation, adverse impact, etc., does not get the same research scrutiny as other selection tools do, although it should.

Retesting

One implementation question that has drawn quite a lot of attention in the past few years is retesting. Researchers have looked at whether retesting (or even simply reapplying) affects performance (Matton et al. 2011), faking (Landers et al. 2010), validity (Lievens et al. 2007), adverse impact (Dunleavy et al. 2008, Schleicher et al. 2010), and applicant reactions (Schleicher et al. 2010). In general, this research suggests a need for caution in how data from retested applicants are considered in research as well as careful consideration of retesting policies by organizations

because of score elevation, accompanying effects on validity, and potential problematic effects on adverse impact. We view this spate of attention to an implementation issue as a positive development; practitioners often lament that there is not a lot of guidance in the selection literature on assessment policy issues and what are truly best practices.

EVALUATING SELECTION SYSTEMS

With regard to the evaluation of selection methods, continued discussion of ways to gather validation evidence and to reduce adverse impact and bias were major topics in the past few years, along with a focus on the attitudes and opinions of selection systems stakeholders, including hiring managers, HR professionals, and applicants.

Validation

One would think that the validation of selection procedures would be a ho-hum, or routine, topic after so many years of research and the establishment of various sets of professional guidelines. However, research guidance on how to conduct criterion validation studies continues to grow (for a review, see Van Iddekinge & Ployhart 2008). As noted previously, we continue to see debate around the usefulness of content validation strategies for supporting selection tool use (see Murphy 2009 and the accompanying set of commentary articles). We also see renewed discussion of the usefulness of synthetic validation methods (see Johnson et al. 2010 and accompanying commentaries). Although there is some practical guidance for conducting transportability studies and for documenting the appropriate use of validity generalization results (Gibson & Caplinger 2007, McDaniel 2007), no empirical work exists on comparisons of practices in this area.

Subgroup Differences and Predictive Bias

Evaluation work also continues on how to reduce subgroup differences (i.e., mean differences between different groups, such as race or sex) and adverse impact (i.e., differences in hiring rates between different subgroups) (Outtz 2010). A closely related topic is how to enhance or change the diversity of the selected workforce (Newman & Lyon 2009). In a recent review, Ryan & Powers (2012) described how many of the choices made in recruitment, in choosing selection system content, and in how selection systems are implemented may make small differences in the diversity of those hired, but when considered in aggregation might have appreciable effects on the adverse impact of hiring processes. However, there still remains a lot of variability in how adverse impact is actually assessed (Biddle & Morris 2011, Cohen et al. 2010, Murphy & Jacobs 2012) and a desire among our practice leaders for further investigation and guidance.

As noted previously, there has been a shift in views regarding predictive bias. Although the received wisdom was once that there is no differential validity with cognitive ability tests, a "revival" of this research (Aguinis et al. 2010, Meade & Tonidandel 2010) has suggested the need to reconsider those conclusions (Berry et al. 2011). One concern with this focus is whether it is a chase of fairly small effects, evidenced only with large sample sizes, and whether renewed research will ultimately result in conclusions different from those of the past. One point raised, however, certainly could change that thinking: whether context moderates the likelihood of finding predictive bias. For example, might the organization's diversity climate, affirmative action policy, selection ratio, and targeted recruiting efforts factor into both who is hired and how they are evaluated once on the job? We expect to see some of this type of work emerging over the next several years.

Stakeholder Goals and Perceptions

We separate our discussion of internal stakeholders from external stakeholders. Internal stakeholders include incumbent employees and managers who must live with the results of the selection hires, selection managers who own responsibility for the design of the selection system, and HR managers who must align the selection system with the organization's strategic goals and economic realities. External stakeholders include applicants affected by the selection system.

Internal stakeholders. Understanding what leads to the adoption and support of selection systems is always a practical concern. Some research effort has been devoted to understanding what influences internal stakeholder perceptions of selection procedures (Furnham 2008, König et al. 2010). Research on utility analysis models, which are intended to provide information on the financial returns from different selection practices, appears to have waned dramatically from its popularity in the 1980s (Cascio & Aguinis 2008). It is not that interest in demonstrating return on investment (ROI) is less important (in fact, given current economic realities, quite the opposite is true). There are still researchers who use traditional utility analysis procedures in an attempt to value employee selection (e.g., Le et al. 2007), but these estimates are usually met with skepticism and concern (e.g., Schmitt 2007). Rather, researchers have focused on better integrating what we know about decision making and persuasion from cognitive and social psychology into methods of presenting cases for changing selection processes. For example, Winkler and colleagues (2010) present a multiattribute supply chain model of ROI that managers may find more persuasive. Similarly, Boudreau (2010) has argued that efforts to convey ROI must be based on existing decision-making models and methods that practicing managers accept and are familiar with. Other research has followed a different approach, adopting methods in strategic human resources, to examine the direct dollar consequences of acquiring higher-quality aggregate KSAOs or using more valid selection practices on unit-level performance (e.g., Ployhart et al. 2009, Van Iddekinge et al. 2009). Examining this issue from a different perspective, Highhouse (2008) noted that managers are frequently resistant to use or be persuaded by validity information of a technical nature, preferring instead to rely on intuition and heuristic judgments. Work on how best to present the case for use of selection tools is likely to continue in the next few years, particularly as new technologies entice stakeholders to use unproven methods and raise concerns over development costs.

External stakeholders. Applicant reactions research has fizzled a bit from a period of strong activity to studies on specific methods less investigated (e.g., reactions to emotional intelligence measures, Iliescu et al. 2012; reactions to credit checks, Kuhn 2012), country-specific investigations (e.g., Saudi Arabia, Anderson et al. 2012; Vietnam, Hoang et al. 2012; Greece, Nikolaou 2011), comparisons of internal and external candidate views (e.g., Giumetti & Sinar 2012), and summaries (Anderson et al. 2010, Truxillo et al. 2009 on explanations). Ryan & Huth (2008) argued that much of the research results in platitudes (e.g., treat people with respect) and therefore has little value for those in practice who are seeking specific guidance on how to improve reactions. We also believe research often focuses on the wrong outcomes, as reactions are unlikely to have much influence on test scores or validity given the typical strong demands and consequences present in a selection system. Rather, we suspect the consequences of reactions are much more subtle, such as spreading negative information about the company, affecting consumer behavior, or harming the company's brand or reputation. These more subtle outcomes may still be quite important from the broader perspective of the business, but they are harder to investigate (i.e., most research simply looks at intentions and not behavior).

Table 1 Forecasted shifts in the selection research base

Globalization
Shift from Western-centric view to multicultural view. Knowledge, skills, abilities, and other characteristics (KSAO) composites, assessments, and selection systems for multicultural competencies (e.g., willingness to work with others from different cultures) become more prominent. In the past, much selection research has been framed within the US legal system. Other countries have very different systems. Further, even within the United States, attitudes are changing, and demographics are changing (e.g., whites are projected to no longer be the majority by 2050). Although adverse impact research has been prominent for the past 50 years (since the US Civil Rights Act of 1964), one would think it might start to decline
Technology
Internet and social media platforms continue to change approaches to assessment and assessment content
Large vendors hold the assessment data on millions of candidates, essentially becoming placement offices where companies can come to them to source viable candidates. This might lead to a revival of placement research, and a very different—and vital—role for vendors
Changes in the nature of work
Increased use of teams and workgroups means a greater focus on relevant interpersonal KSAOs (e.g., teamwork)
Work requires greater knowledge (e.g., growth of professional service industry, high-knowledge jobs), and job knowledge testing becomes more common and sophisticated (e.g., complex simulations)
Working for organizations without homogeneous identities (e.g., limited liability corporations, firms that collaborate in some industries but compete in others, part-time or project-based work) leads to less focus on culture fit and more focus on adaptability
Increased focus on prediction of performance for older workers as those from the baby-boom generation begin to work part time to balance professional and personal preferences
What we need less of:
Demonstrating that small effects "could" occur (i.e., differential prediction/differential validity research)
Continued reestimation of the size of subgroup differences
Predictor-criterion correspondence validity studies (i.e., match the predictor construct with the criterion construct)
Arguing about narrow versus broad predictor usefulness
Demonstrating that applicants can fake
Describing applicant reactions without links to behavioral or objective outcomes
What we need more of:
Theoretically directed research on technologically sophisticated assessments
Psychometrics, scoring, and validity of gamified systems
Prediction of criteria that exist at unit or multiple levels
How to select for company fit, values, and culture
Predicting a broader range of criteria (e.g., adaptive, longitudinal, turnover)
Connecting selection research into the broader talent life cycle
Design, implementation, and evaluation of global selection systems
Greater conceptual specification and investigation of the role of context in selection
Investigation of internal stakeholder perceptions, reactions, judgments, and decision-making processes

CONCLUSIONS

With the scientific field of employee selection turning 100, it is interesting to speculate on what a selection review conducted 100 years from now might look like. Such an exercise is not purely academic but speaks to what is the core of this profession and what we believe is lasting about it. We believe the following questions are worth considering.

- One hundred years from now, will criterion-related validity still be the central question within selection? The predictive hypothesis has remained the core hypothesis in selection

research for the past 100 years (Guion 2011)—could it possibly not be the core question for the next 100 years?

- There is every reasonable expectation that organizations will still need to hire employees in the next century, but does this practical need guarantee the life of employee selection research? For example, could the day come when selection research moves entirely into practice or technical functions? Might selection research be conducted entirely by vendors, in the service of their products and services? Ironically, it is the heavy emphasis on psychometrics that makes selection researchers unique, but this is also the kind of research that has been slowly eroding from the applied psychology journals, being seen as too narrow and too atheoretical. Will the assessment of KSAOs in the future bear any resemblance to the KSAOs of interest today? For example, might selection research evolve into the study of biological or genetic individual differences? If so, would the KSAO measurement methodologies of today be seen as akin to the "introspection" methods from 100 years ago?
- Who will most strongly shape the future of selection research? To date, most of this research has occurred within the United States, and is largely shaped by Western culture. Business already operates in a global economy, and the study of selection will likewise be pulled into this broader context.

Standing on approximately 100 years of employee selection research finds the field in a curious position. While there is great sophistication in selection procedures, methodologies, and practices, the basic question that is of interest to selection researchers—namely, the predictive hypothesis—remains largely unchanged. Some have argued this suggests that selection research has remained invariant, perhaps even indifferent, to changes in the broader world (e.g., Cascio & Aguinis 2008). Others have argued that selection research has shown impressive advancements and remains a vibrant area of research (e.g., Sackett & Lievens 2008). We offer yet a third perspective: that traditional selection research will remain active and engaging, but the broader challenges facing the field will push selection research into exciting new areas and in turn attract a broader array of scientist-practitioners. **Table 1** offers a glimpse of some of those directions and what are likely to be areas of emphasis and de-emphasis in research. Selection research will expand to consider a broader range of criteria (within and across levels), in different cultures and contexts, and in an ever-expanding array of practical applications. Viva selection!

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

Aguinis H, Culpepper SA, Pierce CA. 2010. Revival of test bias research in pre-employment testing. *J. Appl. Psychol.* 95:648–80

Anderson N, Ahmed S, Ana CC. 2012. Applicant reactions in Saudi Arabia: organizational attractiveness and core-self evaluation. *Int. J. Sel. Assess.* 20(2):197–208

Anderson N, Salgado JF, Hulsheger UR. 2010. Applicant reactions in selection: comprehensive meta-analysis into reaction generalization versus situational specificity. *Int. J. Sel. Assess.* 18:291–304

Arthur W, Glaze RM, Villado AJ, Taylor JE. 2010. The magnitude and extent of cheating and response distortion effects on unproctored internet-based tests of cognitive ability and personality. *Int. J. Sel. Assess.* 18:1–16

Arthur W, Villado AJ. 2008. The importance of distinguishing between constructs and methods when comparing predictors in personnel selection research and practice. *J. Appl. Psychol.* 93:435–42

Baker E, Demps J. 2009. Videoconferencing as a tool for recruiting and interviewing. *J. Bus. Econ. Res.* 7(10):9–14

Banks GC, Kepes S, McDaniel MA. 2012. Publication bias: a call for improved meta-analytic practice in the organizational sciences. *Int. J. Sel. Assess.* 20:182–96

Barrick MR, Shaffer JA, DeGrassi SW. 2009. What you see may not be what you get: a meta-analysis of the relationship between self-presentation tactics and ratings of interview and job performance. *J. Appl. Psychol.* 94:1394–411

Barrick MR, Swider B, Stewart GL. 2010. Initial evaluations in the interview: relationships with subsequent interviewer evaluations and employment offers. *J. Appl. Psychol.* 95:1037–46

Barrick MR, Zimmerman RD. 2009. Hiring for retention and performance. *Hum. Resour. Manag.* 48:183–206

Bartram D. 2008. Global norms: towards some guidelines for aggregating personality norms across countries. *Int. J. Test.* 8(4):315–33

Bartram D. 2009. The international test commission guidelines on computer-based and internet-delivered testing. *Ind. Organ. Psychol.: Perspect. Sci. Pract.* 2(1):11–13

Bartram D. 2013. Scalar equivalence of OPQ32: Big Five profiles of 31 countries. *J. Cross-Cult. Psychol.* 44(1):61–83

Beaty JC, Nye CD, Borneman MJ, Kantrowitz TM, Drasgow F, Grauer E. 2011. Proctored versus unproctored Internet tests: Are unproctored noncognitive tests as predictive of job performance? *Intl. J. Sel. Assess.* 19:1–10

Bernerth JB, Taylor SG, Walker HJ, Whitman DS. 2012. An empirical investigation of dispositional antecedents and performance-related outcomes of credit scores. *J. Appl. Psychol.* 97(2):469–78

Berry CM, Clark MA, McClure TK. 2011. Racial/ethnic differences in the criterion-related validity of cognitive ability tests: a qualitative and quantitative review. *J. Appl. Psychol.* 96(5):881–906

Berry CM, Sackett PR. 2009. Faking in personnel selection: tradeoffs in performance versus fairness resulting from two cut-score strategies. *Pers. Psychol.* 62(4):835–63

Biddle DA, Morris SB. 2011. Using Lancaster's mid-P correction to the Fisher's exact test for adverse impact analyses. *J. Appl. Psychol.* 96(5):956–65

Binning JF, Lebreton JM. 2009. Coherent conceptualization is useful for many things, and understanding validity is one of them. *Ind. Organ. Psychol.: Perspect. Sci. Pract.* 2(4):486–92

Bledow R, Frese M. 2009. A situational judgment test of personal initiative and its relationship to performance. *Pers. Psychol.* 62:229–58

Blickle G, Momm TS, Kramer J, Mierke J, Liu Y, Ferris GR. 2009. Construct and criterion-related validation of a measure of emotional reasoning skills: a two-study investigation. *Int. J. Sel. Assess.* 17:101–18

Bobko P, Roth PL. 2013. Reviewing, categorizing, and analyzing the literature on black–white mean differences for predictors of job performance: verifying some perceptions and updating/correcting others. *Pers. Psychol.* 66:91–126

Boudreau JW. 2010. *Retooling HR*. Boston, MA: Harv. Bus. Sch. Publ.

Breaugh JA. 2009. The use of biodata for employee selection: past research and future directions. *Hum. Resour. Manag. Rev.* 19:219–31

Breuker J. 2013. A cognitive science perspective on knowledge acquisition. *Int. J. Hum. Comput. Stud.* 71(2):177–83

Brown V, Vaughn E. 2011. The writing on the (Facebook) wall: the use of social networking sites in hiring decisions. *J. Bus. Psychol.* 10:11–21

Brummel BJ, Rupp DE, Spain SM. 2009. Constructing parallel simulation exercises for assessment centers and other forms of behavioral assessment. *Pers. Psychol.* 62(1):135–70

Caligiuri P, Paul KB. 2010. *Selection in Multinational Organizations*. New York: Routledge/Taylor & Francis

Caligiuri P, Tarique I, Jacobs R. 2009. Selection for international assignments. *Hum. Resour. Manag. Rev.* 19:251–62

Campbell JP. 1990. *Modeling the Performance Prediction Problem in Industrial and Organizational Psychology*. Palo Alto, CA: Consult. Psychol. Press

Campion MA, Fink AA, Ruggeberg BJ, Carr L, Phillips GM, Odman RB. 2011. Doing competencies well: best practices in competency modeling. *Pers. Psychol.* 64(1):225–62

Carey T, Herst D, Chan W. 2010. Global selection: selection in international contexts. In *Going Global: Practical Applications and Recommendations for HR and OD Professionals in the Global Workplace*, ed. K Lundby, pp. 143–74. San Francisco, CA: Jossey-Bass

Cascio WF, Aguinis H. 2008. Research in industrial and organizational psychology from 1963 to 2007: changes, choices, and trends. *J. Appl. Psychol.* 93:1062–81

Chapman DS, Zweig DI. 2005. Developing a nomological network for interview structure: antecedents and consequences of the structured selection interview. *Pers. Psychol.* 58(3):673–702

Chen YC, Tsai WC, Hu C. 2008. The influences of interviewer-related and situational factors on interviewer reactions to high structured job interviews. *Int. J. Hum. Resour. Manag.* 19:1056–71

Cherniss C. 2010. Emotional intelligence: towards clarification of a concept. *Ind. Organ. Psychol.: Perspect. Sci. Pract.* 3:110–26

Chiaburu DS, Oh I, Berry CM, Li N, Gardner RG. 2011. The five-factor model of personality traits and organizational citizenship behaviors: a meta-analysis. *J. Appl. Psychol.* 96(6):1140–66

Christian MS, Edwards BD, Bradley JC. 2010. Situational judgment tests: constructs assessed and a meta-analysis of their criterion-related validities. *Pers. Psychol.* 63:83–117

Christiansen ND, Janovics JE, Siers BP. 2010. Emotional intelligence in selection contexts: measurement method, criterion-related validity, and vulnerability to response distortion. *Int. J. Sel. Assess.* 18:87–101

Cleveland JN, Colella A. 2010. *Criterion Validity and Criterion Deficiency: What We Measure Well and What We Ignore*. New York: Routledge/Taylor & Francis

Coff R, Kryscynski D. 2011. Drilling for micro-foundations of human capital–based competitive advantages. *J. Manag.* 37:1429–43

Cohen DB, Aamodt MG, Dunleavy EM. 2010. *Technical Advisory Committee Report on Best Practices in Adverse Impact Analyses*. Washington, DC: Cent. Corp. Excell.

Colom R, Martínez-Molina A, Shih PC, Santacreu J. 2010. Intelligence, working memory, and multitasking performance. *Intelligence* 38(6):543–51

Compton P. 2013. Situated cognition and knowledge acquisition research. *Int. J. Hum. Comput. Stud.* 71(2):184–90

Converse PD, Oswald FL, Imus A, Hedricks C, Roy R, Butera H. 2008. Comparing personality test formats and warnings: effects on criterion-related validity and test-taker reactions. *Int. J. Sel. Assess.* 16:155–69

Converse PD, Pathak J, Quist J, Merbedone M, Gotlib T, Kostic E. 2010. Statement desirability ratings in forced-choice personality measure development: implications for reducing score inflation and providing trait-level information. *Hum. Perform.* 23(4):323–42

Crook TR, Todd SY, Combs JG, Woehr DJ, Ketchen DJ Jr. 2011. Does human capital matter? A meta-analysis of the relationship between human capital and firm performance. *J. Appl. Psychol.* 96(3):443–56

Cucina JM, Caputo PM, Thibodeaux HF, Maclane CN. 2012. Unlocking the key to biodata scoring: a comparison of empirical, rational, and hybrid approaches at different sample sizes. *Pers. Psychol.* 65(2):385–428

Davison HK, Maraist CC, Bing MN. 2011. Friend or foe? The promise and pitfalls of using social networking sites for HR activities. *J. Bus. Psychol.* 26:153–59

Dean MA, Roth PL, Bobko P. 2008. Ethnic and gender subgroup differences in assessment center ratings: a meta-analysis. *J. Appl. Psychol.* 93(3):685–91

DeCorte W, Sackett P, Lievens F. 2010. Selecting predictor subsets: considering validity and adverse impact. *Int. J. Sel. Assess.* 18(3):260–70

DeCorte W, Sackett PR, Lievens F. 2011. Designing pareto-optimal selection systems: formalizing the decisions required for selection system development. *J. Appl. Psychol.* 96(5):907–26

Derous E, Nguyen H, Ryan AM. 2009. Hiring discrimination against Arab minorities: interactions between prejudice and job characteristics. *Hum. Perform.* 22(4):297–320

Derous E, Ryan AM, Nguyen HD. 2012. Multiple categorization in resume screening: examining effects on hiring discrimination against Arab applicants in field and lab settings. *J. Organ. Behav.* 33(4):544–70

Dilchert S, Ones DS. 2009. Assessment center dimensions: individual differences correlates and meta-analytic incremental validity. *Int. J. Sel. Assess.* 17:254–70

Dorsey DW, Cortina JM, Luchman J. 2010. *Adaptive and Citizenship-Related Behaviors at Work*. New York: Routledge/Taylor & Francis

Drover W, Franczak J, Beltramini RF. 2012. A 30-year historical examination of ethical concerns regarding business ethics: Who's concerned? *J. Bus. Ethics* 111(4):431–38

Dunleavy EM, Mueller LM, Buonasera AK, Kuang DC, Dunleavy DG. 2008. On the consequences of frequent applicants in adverse impact analyses: a demonstration study. *Int. J. Sel. Assess.* 16:334–44

Eisner S. 2010. E-employment? College grad career building in a changing and electronic age. *Am. J. Bus. Educ.* 3(7): 25–40

Ellingson JA, McFarland LA. 2011. Understanding faking behavior through the lens of motivation: an application of VIE theory. *Hum. Perform.* 24:322–37

Fan J, Gao D, Carroll SA, Lopez FJ, Tian TS, Meng H. 2012. Testing the efficacy of a new procedure for reducing faking on personality tests within selection contexts. *J. Appl. Psychol.* 97(4):866–80

Farr JL, Tippins NT. 2010. *Handbook of Employee Selection: An Introduction and Overview*. New York: Routledge/Taylor & Francis

Finch DM, Edwards BD, Wallace JC. 2009. Multistage selection strategies: simulating the effects on adverse impact and expected performance for various predictor combinations. *J. Appl. Psychol.* 94:318–40

Foldes HJ, Duehr EE, Ones DS. 2008. Group differences in personality: meta-analyses comparing five U.S. racial groups. *Pers. Psychol.* 61(3):579–616

Furnham A. 2008. HR professionals' beliefs about, and knowledge of, assessment techniques and psychometric tests. *Int. J. Sel. Assess.* 16(3):300–5

Gibbons AM, Rupp DE. 2009. Dimension consistency as an individual difference: a new (old) perspective on the assessment center construct validity debate. *J. Manag.* 35:1154–80

Gibson WM, Caplinger JA. 2007. Transportation of validation results. In *Alternative Validation Strategies: Developing New and Leveraging Existing Validity Evidence*, ed. SM McPhail, pp. 29–81. New York: Wiley

Giumetti GW, Sinar EF. 2012. Don't you know me well enough yet? Comparing reactions of internal and external candidates to employment testing. *Int. J. Sel. Assess.* 20(2):139–48

Griffith RL, Peterson MH. 2011. One piece at a time: the puzzle of applicant faking and a call for theory. *Hum. Perform.* 24:291–301

Guion RM 2011. *Assessment, Measurement, and Prediction for Personnel Decisions*. New York: Routledge

Gutl C, Lankmayr K, Weinhofer J, Hofler M. 2011. Enhanced Automatic Question Creator—EAQC: concept, development and evaluation of an automatic test item creation tool to foster modern e-education. *Electron. J. e-Learn.* 9(1):23–38

Harms PD, Crede M. 2010. Remaining issues in emotional intelligence research: construct overlap, method artifacts and lack of incremental validity. *Ind. Organ. Psychol.: Perspect. Sci. Pract.* 3:154–58

Hattrup K. 2012. Using composite predictors in personnel selection. See Schmitt 2012, pp. 297–319

Highhouse S. 2008. Stubborn reliance on intuition and subjectivity in employee selection. *Ind. Organ. Psychol.: Perspect. Sci. Pract.* 1(3):333–42

Hoang TG, Truxillo DM, Erdogan B, Bauer TN. 2012. Cross-cultural examination of applicant reactions to selection methods: United States and Vietnam. *Int. J. Sel. Assess.* 20(2):209–19

Hoffman BJ, Meade A. 2012. Alternate approaches to understanding the psychometric properties of assessment centers: an analysis of the structure and equivalence of exercise ratings. *Int. J. Sel. Assess.* 20(1):82–97

Hoffman BJ, Melchers KG, Blair CA, Kleinmann M, Ladd RT. 2011. Exercises and dimensions are the currency of assessment centers. *Pers. Psychol.* 64(2):351–95

Hough L, Dilchert S. 2010. Personality: its measurement and validity for employee selection. In *Handbook of Employee Selection*, ed. J Farr, N Tippins, pp. 299–319. New York: Routledge/Taylor & Francis

Huffcutt AI. 2011. An empirical review of the employment interview construct literature. *Int. J. Sel. Assess.* 19:62–81

Iliescu D, Ilie A, Ispas D, Ion A. 2012. Emotional intelligence in personnel selection: applicant reactions, criterion, and incremental validity. *Int. J. Sel. Assess.* 20(3):347–58

Inceoglu I, Bartram D. 2012. Global leadership: the myth of multicultural competency. *Ind. Organ. Psychol.: Perspect. Sci. Pract.* 5(2):216–18

Indulska M, Dirk SH, Recker J. 2012. Quantitative approaches to content analysis: identifying conceptual drift across publication outlets. *Eur. J. Inf. Syst.* 21(1):49–69

Jansen A, Melchers KG, Lievens F, Kleinmann M, Brändli M, et al. 2012. Situation assessment as an ignored factor in the behavioral consistency paradigm underlying the validity of personnel selection procedures. *J. Appl. Psychol.* 96:326–41

Johnson JW, Steel P, Scherbaum CA, Hoffman CC, Jeanneret PR, Foster J. 2010. Validation is like motor oil: Synthetic is better. *Ind. Organ. Psychol.* 3:305–28

Jones RG, Born M. 2008. Assessor constructs in use as the missing component in validation of assessment center dimensions: a critique and directions for research. *Int. J. Sel. Assess.* 16:229–38

Joseph DL, Newman DA. 2010. Emotional intelligence: an integrative meta-analysis and cascading model. *J. Appl. Psychol.* 95(1):54–78

Kantrowitz TM, Dawson CR, Fetzer MS. 2011. Computer adaptive testing (CAT): a faster, smarter, and more secure approach to pre-employment testing. *J. Bus. Psychol.* 26(2):227–32

Kantrowitz TM, Grelle DM, Beaty JC, Wolf MB. 2012. Time is money: polychronicity as a predictor of performance across job levels. *Hum. Perform.* 25(2):114–37

Kehoe JF. 2008. Commentary on pareto-optimality as a rationale for adverse impact reduction: What would organizations do? *Int. J. Sel. Assess.* 16(3):195–200

Kepes S, Banks GC, McDaniel MA, Whetzel DL. 2012. Publication bias in the organizational sciences. *Organ. Res. Methods* 15:624–62

Klehe UC, Konig CJ, Richter GM, Kleinmann M, Melchers KG. 2008. Transparency in structured interviews: consequences for construct and criterion-related validity. *Hum. Perform.* 21:107–37

Klehe UC, Latham GP. 2005. The predictive and incremental validity of the situational and patterned behavior description interviews for team playing behavior. *Int. J. Sel. Assess.* 13(2):108–15

Kleinmann M, Klehe UC. 2011. Selling oneself: construct and criterion-related validity of impression management in structured interviews. *Hum. Perform.* 24:29–46

Kluemper DH, Rosen PA. 2009. Future employment selection methods: evaluating social networking web sites. *J. Manag. Psychol.* 24(6):567–80

Komar S, Brown DJ, Komar JA, Robie C. 2008. Faking and the validity of conscientiousness: a Monte Carlo investigation. *J. Appl. Psychol.* 93:140–54

König CJ, Klehe U, Berchtold M, Kleinmann M. 2010. Reasons for being selective when choosing personnel selection procedures. *Int. J. Sel. Assess.* 18:17–27

Kuhn K. 2012. The controversy over credit checks in selection: using public discourse to gain insight into divergent beliefs, concerns, and experiences. *J. Manag. Inq.* 21(3):331–47

Kuncel NR. 2008. Some new (and old) suggestions for improving personnel selection. *Ind. Organ. Psychol.: Perspect. Sci. Pract.* 1(3):343–46

LaHuis DM, Copeland D. 2009. Investigating faking using a multilevel logistic regression approach to measuring person fit. *Organ. Res. Methods* 12(2):296–319

Landers RN, Sackett PR, Tuzinski KA. 2011. Retesting after initial failure, coaching rumors, and warnings against faking in online personality measures for selection. *J. Appl. Psychol.* 96(1):202–10

Lang JWB, Kersting M, Hulscheger UR, Lang J. 2010. General mental ability, narrower cognitive abilities, and job performance: the perspective of the nested-factors model of cognitive abilities. *Pers. Psychol.* 63:595–640

Le H, Oh IS, Robbins SB, Ilies R, Holland E, Westrick P. 2011. Too much of a good thing: curvilinear relationships between personality traits and job performance. *J. Appl. Psychol.* 96(1):113–33

Le H, Oh IS, Shaffer JA, Schmidt FL. 2007. Implications of methodological advances for the practice of personnel selection: how practitioners benefit from recent developments in meta-analysis. *Acad. Manag. Perspect.* 21(3):6–15

Levashina J, Morgeson FP, Campion MA. 2012. Tell me some more: exploring how verbal ability and item verifiability influence responses to biodata questions in a high-stakes selection context. *Pers. Psychol.* 65(2):359–83

Lievens F, Klehe U, Libbrecht N. 2011. Applicant versus employee scores on self-report emotional intelligence measures. *J. Pers. Psychol.* 10(2):89–95

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. *J. Appl. Psychol.* 96(5):927–40

Lievens F, Peeters H, Schollaert E. 2008. Situational judgment tests: a review of recent research. *Pers. Rev.* 37:426–41

Lievens F, Reeve CL, Heggestad ED. 2007. An examination of psychometric bias due to retesting on cognitive ability tests in selection settings. *J. Appl. Psychol.* 92(6):1672–82

Lievens F, Sackett PR. 2012. The validity of interpersonal skills assessment via situational judgment tests for predicting academic success and job performance. *J. Appl. Psychol.* 97(2):460–68

Lievens F, Sackett PR, Buyse T. 2009. The effects of response instructions on situational judgment test performance and validity in a high-stakes context. *J. Appl. Psychol.* 94:1095–101

Lievens F, Van Keer E, Volckaert E. 2010. Gathering behavioral samples through a computerized and standardized assessment center exercise: Yes, it is possible. *J. Pers. Psychol.* 9(2):94–98

Luecht RM. 2000. *Implementing the computer-adaptive sequential testing (CAST) framework to mass produce high quality computer-adaptive and mastery tests*. Paper presented at Annu. Meet. Natl. Counc. Meas. Educ., New Orleans, La., April 25–27

Macan T. 2009. The employment interview: a review of current studies and directions for future research. *Hum. Resour. Manag. Rev.* 19:203–18

MacKenzie WI, Ployhart RE, Weekley JA, Ehlers C. 2010. Contextual effects on SJT responses: an examination of construct validity and mean differences across applicant and incumbent contexts. *Hum. Perform.* 23:1–21

Maltarich MA, Nyberg AJ, Reilly G. 2010. A conceptual and empirical analysis of the cognitive ability–voluntary turnover relationship. *J. Appl. Psychol.* 95:1058–70

Marcus B. 2009. "Faking" from the applicant's perspective: a theory of self-presentation in personnel selection settings. *Int. J. Sel. Assess.* 17:417–30

Matton N, Vautier S, Raufaste E. 2011. Test-specificity of the advantage of retaking cognitive ability tests. *Int. J. Sel. Assess.* 19:11–17

Maurer SD, Cook DP. 2011. Using company web sites to e-recruit qualified applicants: a job marketing based review of theory-based research. *Comput. Hum. Behav.* 27(1):106–17

McDaniel MA. 2007. Validity generalization as a test validation approach. In *Alternative Validation Strategies: Developing New and Leveraging Existing Validity Evidence*, ed. SM McPhail, pp. 159–80. New York: Wiley

McDaniel MA, Psotka J, Legree PJ, Yost AP, Weekley JA. 2011. Toward an understanding of situational judgment item validity and group differences. *J. Appl. Psychol.* 96:327–36

McNelly T, Ruggeberg BJ, Hall CR. 2011. Web-based management simulations: technology-enhanced assessment for executive-level selection and development. In *Technology-Enhanced Assessment of Talent*, ed. NT Tippins, S Adler, pp. 253–66. San Francisco, CA: Jossey-Bass

Meade AW, Tonidandel S. 2010. Not seeing clearly with Cleary: what test bias analyses do and do not tell us. *Ind. Organ. Psychol.: Perspect. Sci. Pract.* 3(2):192–205

Melchers KG, Kleinmann M, Prinz MA. 2010. Do assessors have too much on their plates? The effects of simultaneously rating multiple assessment center candidates on rating quality. *Int. J. Sel. Assess.* 18(3): 329–41

Melchers KG, Lienhardt N, von Aarburg M, Kleinmann M. 2011. Is more structure really better? A comparison of frame-of-reference training and descriptively anchored rating scales to improve interviewers' rating quality. *Pers. Psychol.* 64:53–87

Meriac JP, Hoffman BJ, Woehr DJ, Fleisher MS. 2008. Further evidence for the validity of assessment center dimensions: a meta-analysis of the incremental criterion-related validity of dimension ratings. *J. Appl. Psychol.* 93(5):1042–52

Meyer JP, Hecht TD, Gill H, Toplonytsky L. 2010. Person-organization (culture) fit and employee commitment under conditions of organizational change: a longitudinal study. *J. Vocat. Behav.* 76(3):458–73

Motowidlo SJ, Beier ME. 2010. Differentiating specific job knowledge from implicit trait policies in procedural knowledge measured by a situational judgment test. *J. Appl. Psychol.* 95:321–33

Mount MK, Oh I, Burns M. 2008. Incremental validity of perceptual speed and accuracy over general mental ability. *Pers. Psychol.* 61(1):113–39

Mumford TV, Van Iddekinge CH, Morgeson FP, Campion MA. 2008. The team role test: development and validation of a team role knowledge situational judgment test. *J. Appl. Psychol.* 93:250–67

Munsterberg H. 1913. *Psychology and Industrial Efficiency*. Boston, MA/New York: Houghton Mifflin
Murphy KR. 2009. Content validation is useful for many things, but validity isn't one of them. *Ind. Organ. Psychol.: Perspect. Sci. Pract.* 2:465–68
Murphy KR, Dzieweczynski JL, Zhang Y. 2009. Positive manifold limits the relevance of content-matching strategies for validating selection test batteries. *J. Appl. Psychol.* 94(4):1018–31
Murphy KR, Jacobs RR. 2012. Using effect size measures to reform the determination of adverse impact in equal employment litigation. *Psychol. Public Policy Law* 18(3):477–99
Ndoye A, Ritzhaupt AD, Parker MA. 2012. Use of ePortfolios in K-12 teacher hiring in North Carolina: perspectives of school principals. *Int. J. Educ. Policy Leadersh.* 7(4):1–10
Newman DA, Joseph DL, MacCann C. 2010. Emotional intelligence and job performance: the importance of emotion regulation and emotional labor context. *Ind. Organ. Psychol.: Perspect. Sci. Pract.* 3:159–64
Newman DA, Lyon JS. 2009. Recruitment efforts to reduce adverse impact: targeted recruiting for personality, cognitive ability, and diversity. *J. Appl. Psychol.* 94:298–317
Nikolaou I. 2011. Core processes and applicant reactions to the employment interview: an exploratory study in Greece. *Int. J. Hum. Resour. Manag.* 22(10):2185–201
O'Boyle EH, Forsyth DR, Banks GC, McDaniel MA. 2012. A meta-analysis of the dark triad and work behavior: a social exchange perspective. *J. Appl. Psychol.* 97(3):557–79
O'Brien J. 2008. Interviewer resistance to structure. *Ind. Organ. Psychol.: Perspect. Sci. Pract.* 1(3):367–69
Ones DS, Dilchert S, Viswesvaran C, Salgado JF. 2010. Cognitive abilities. In *Handbook of Employee Selection*, ed. JL Farr, NT Tippins, pp. 255–76. Mahwah, NJ: Erlbaum
Ones DS, Viswesvaran C, Schmidt FL. 2012. Integrity tests predict counterproductive work behaviors and job performance well: comment on Van Iddekinge, Roth, Raymark, and Odle-Dusseau 2012. *J. Appl. Psychol.* 97(3):537–42
Oostrom JK, Bos-Broekema L, Serlie AW, Born MP, Van der Molen HT. 2012. A field study of pretest and posttest reactions to a paper-and-pencil and a computerized in-basket exercise. *Hum. Perform.* 25(2): 95–113
Ordanini A, Silvestri G. 2008. Recruitment and selection services: efficiency and competitive reasons in the outsourcing of HR practices. *Int. J. Hum. Resour. Manag.* 19:372–91
Oswald FL, Hough L. 2012. I-O 2.0 from intelligence 1.5: staying (just) behind the cutting edge of intelligence theories. *Ind. Organ. Psychol.* 5:172–75
Outtz JL, ed. 2010. *Adverse Impact: Implications for Organizational Staffing and High Stakes Selection*. New York: Routledge/Taylor & Francis
Ployhart RE, Holtz BC. 2008. The diversity-validity dilemma: strategies for reducing racioethnic and sex subgroup differences and adverse impact in selection. *Pers. Psychol.* 61(1):153–72
Ployhart RE, Moliterno TP. 2011. Emergence of the human capital resource: a multilevel model. *Acad. Manag. Rev.* 36:127–50
Ployhart RE, Van Iddekinge C, MacKenzie W. 2011. Acquiring and developing human capital in service contexts: the interconnectedness of human capital resources. *Acad. Manag. J.* 54:353–68
Ployhart RE, Weekley JA, Ramsey J. 2009. The consequences of human resource stocks and flows: a longitudinal examination of unit service orientation and unit effectiveness. *Acad. Manag. J.* 52:996–1015
Postlethwaite BE, Giluk TL, Schmidt FL. 2012. I-O psychologists and intelligence research: active, aware, and applied. *Ind. Organ. Psychol.* 5:186–88
Potosky D. 2008. A conceptual framework for the role of the administration medium in the personnel assessment process. *Acad. Manag. Rev.* 33(3):629–48
PriceWaterhouseCoopers. 2010. *Talent Mobility 2020: The Next Generation of International Assignments*. **http://www.pwc.com/gx/en/managing-tomorrows-people/future-of-work/pdf/talent-mobility-2020.pdf**
Putka DJ, Hoffman BJ. 2013. Clarifying the contribution of assessee-, dimension-, exercise-, and assessor-related effects to reliable and unreliable variance in assessment center ratings. *J. Appl. Psychol.* 98(1):114–33
Pyburn KM, Ployhart RE, Kravitz DA. 2008. The diversity-validity dilemma: overview and legal context. *Pers. Psychol.* 61(1):143–51
Riggio RE. 2010. Before emotional intelligence: research on nonverbal, emotional and social competencies. *Ind. Organ. Psychol.: Perspect. Sci. Pract.* 3:178–82

Ryan AM, Boyce AS, Ghumman S, Jundt D, Schmidt G, Gibby R. 2009. Going global: cultural values and perceptions of selection procedures. *Appl. Psychol.* 58(4):520–56

Ryan AM, Huth M. 2008. Not much more than platitudes? A critical look at the utility of applicant reactions research. *Hum. Resour. Manag. Rev.* 18(3):119–32

Ryan AM, Powers C. 2012. Workplace diversity. See Schmitt 2012, pp. 814–31

Ryan AM, Tippins N. 2009. *Designing and Implementing Global Selection Systems*. West Sussex, UK: Wiley-Blackwell

Sackett PR, Lievens F. 2008. Personnel selection. *Annu. Rev. Psychol.* 59:419–50

Sackett PR, Schmitt N. 2012. On reconciling conflicting meta-analytic findings regarding integrity test validity. *J. Appl. Psychol.* 97:550–56

Sanchez J, Levine E. 2012. The rise and fall of job analysis and the future of work analysis. *Annu. Rev. Psychol.* 63:397–425

Scherbaum C, Goldstein H, Yusko K, Ryan R, Hanges PJ. 2012. Intelligence 2.0: reestablishing a research program on *g* in I-O psychology. *Ind. Organ. Psychol.* 5:128–48

Schippman JS. 2010. Competencies, job analysis, and the next generation of modeling. In *Handbook of Workplace Assessment*, ed. JC Scott, DH Reynolds, pp. 197–231. San Francisco, CA: Jossey-Bass

Schleicher DJ, Van Iddekinge CH, Morgeson FP, Campion MA. 2010. If at first you don't succeed, try, try again: understanding race, age, and gender differences in retesting score improvement. *J. Appl. Psychol.* 95(4):603–17

Schmitt N. 2007. The value of personnel selection: reflections on some remarkable claims. *Acad. Manag. Perspect.* 21:19–23

Schmitt N. 2012. *The Oxford Handbook of Personnel Assessment and Selection*. New York: Oxford Univ. Press

Schollaert E, Lievens F. 2012. Building situational stimuli in assessment center exercises: Do specific exercise instructions and role-player prompts increase the observability of behavior? *Hum. Perform.* 25(3):255–71

Scott JC, Lezotte DV. 2012. Web-based assessments. See Schmitt 2012, pp. 485–513

Silzer R, Jeanneret R. 2011. Individual psychological assessment: a practice and science in search of common ground. *Ind. Organ. Psychol.* 4:270–96

Sojka JZ, Deeter-Schmelz D. 2008. Need for cognition and affective orientation as predictors of sales performance: an investigation of main and interaction effects. *J. Bus. Psychol.* 22(3):179–90

Stark S, Chernyshenko OS, Drasgow F, White LA. 2012. Adaptive testing with multidimensional pairwise preference items: improving the efficiency of personality and other noncognitive assessments. *Organ. Res. Methods* 15(3):463–87

Stewart GL, Darnold T, Barrick MR, Dustin SD. 2008. Exploring the handshake in employment interviews. *J. Appl. Psychol.* 93:1139–46

Swider BW, Barrick MR, Harris TB, Stoverink AC. 2011. Managing and creating an image in the interview: the role of interviewee initial impressions. *J. Appl. Psychol.* 96:1275–88

Tett RP, Simonet DV. 2011. Faking in personality assessment: a "multisaturation" perspective on faking as performance. *Hum. Perform.* 24(4):302–21

Thornton GC. 2009. Evidence of content matching is evidence of validity. *Ind. Organ. Psychol.: Perspect. Sci. Pract.* 2(4):469–74

Thornton GC, Gibbons AM. 2009. Validity of assessment centers for personnel selection. *Hum. Resour. Manag. Rev.* 19:169–87

Tippins NT. 2009. Internet alternatives to traditional proctored testing: Where are we now? *Ind. Organ. Psychol.: Perspect. Sci. Pract.* 2(1):2–10

Tippins NT, Adler S. 2011 *Technology-Enhanced Assessment of Talent.* San Francisco, CA: Jossey-Bass

Truxillo DM, Bodner TE, Bertolino M, Bauer TN, Yonce CA. 2009. Effects of explanations on applicant reactions: a meta-analytic review. *Int. J. Sel. Assess.* 17(4):346–61

Tsai W, Chi N, Huang T, Hsu A. 2011. The effects of applicant résumé contents on recruiters' hiring recommendations: the mediating roles of recruiter fit perceptions. *Appl. Psychol.* 60(2):231–54

van Hooft EAJ, Born MP. 2012. Intentional response distortion on personality tests: using eye-tracking to understand response processes when faking. *J. Appl. Psychol.* 97(2):301–16

Van Iddekinge CH, Ferris GR, Perrewe PL, Perryman AA, Blass FR, Heetderks TD. 2009. Effects of selection and training on unit-level performance over time: a latent growth modeling approach. *J. Appl. Psychol.* 94:829–43

Van Iddekinge CH, Ployhart RE. 2008. Developments in the criterion-related validation of selection procedures: a critical review and recommendations for practice. *Pers. Psychol.* 61(4):871–925

Van Iddekinge CH, Putka DJ, Campbell JP. 2011a. Reconsidering vocational interests for personnel selection: the validity of an interest-based selection test in relation to job knowledge, job performance, and continuance intentions. *J. Appl. Psychol.* 96(1):13–33

Van Iddekinge CH, Roth PL, Putka D, Lanivich SE. 2011b. Are you interested? A meta-analysis of relations between vocational interests and employee performance and turnover. *J. Appl. Psychol.* 96:1167–94

Van Iddekinge CH, Roth PL, Raymark PH, Odle-Dusseau H. 2012. The criterion-related validity of integrity tests: an updated meta-analysis. *J. Appl. Psychol.* 97(3):499–530

Westring AJF, Oswald FL, Schmitt N, Drzakowski S, Imus A, et al. 2009. Estimating trait and situational variance in a situational judgment test. *Hum. Perform.* 22:44–63

Whetzel DL, McDaniel MA. 2009. Situational judgment tests: an overview of current research. *Hum. Resour. Manag. Rev.* 19:188–202

Whetzel DL, McDaniel MA, Nguyen NT. 2008. Subgroup differences in situational judgment test performance: a meta-analysis. *Hum. Perform.* 21(3):291–309

Winkler S, König CJ, Kleinmann M. 2010. Single attribute utility analysis may be futile, but this can't be the end of the story: causal chain analysis as an alternative. *Pers. Psychol.* 63:1041–65

Wright PM, McMahan GC. 2011. Exploring human capital: putting human back into strategic human resource management. *Hum. Resour. Manag. J.* 21:93–104

Zickar MJ, Lake CJ. 2011. Practice agenda: innovative uses of technology-enhanced assessment. In *Technology-Enhanced Assessment of Talent*, ed. NT Tippins, S Adler, pp. 394–417. San Francisco, CA: Jossey-Bass

Zualkernan IA, Raza A, Karim A. 2012. Curriculum-guided crowd sourcing of assessments in a developing country. *J. Educ. Technol. Soc.* 15(4):14–26

Annual Review of Psychology

Volume 65, 2014

Contents