 SHRM Foundation's
Effective Practice Guidelines Series

Choosing Effective Talent Assessments to Strengthen Your Organization



SHRM[®]
SOCIETY FOR HUMAN
RESOURCE MANAGEMENT

SHRM[®]
Foundation
Shaping the future of HR

Choosing Effective Talent Assessments to Strengthen Your Organization

This publication is designed to provide accurate and authoritative information regarding the subject matter covered. Neither the publisher nor the authors are engaged in rendering legal or other professional service. If legal advice or other expert assistance is required, the services of a competent, licensed professional should be sought. Any federal and state laws discussed in this report are subject to frequent revision and interpretation by amendments or judicial revisions that may significantly affect employer or employee rights and obligations. Readers are encouraged to seek legal counsel regarding specific policies and practices in their organizations.

This book is published by the SHRM Foundation, an affiliate of the Society for Human Resource Management (SHRM®). The interpretations, conclusions and recommendations in this report are those of the authors and do not necessarily represent those of the SHRM Foundation.

©2016 SHRM Foundation. All rights reserved. Printed in the United States of America.

This publication may not be reproduced, stored in a retrieval system or transmitted in whole or in part, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the SHRM Foundation, 1800 Duke Street, Alexandria, VA 22314.

Selection of report topics, treatment of issues, interpretation and other editorial decisions for the Effective Practice Guidelines series are handled by SHRM Foundation staff and the report authors. Report sponsors may review the content prior to publication and provide input along with other reviewers; however, the SHRM Foundation retains final editorial control over the reports. Editorial decisions are based solely on the defined scope of the report, the accuracy of the information and the value it will provide to the readers.

The SHRM Foundation does not explicitly or by implication endorse or make any representations or warranties of any kind regarding its sponsors or the products, services or claims made by its sponsors. The SHRM Foundation does not assume any responsibility or liability for the acts, omissions, products or services offered by its sponsors.

For more information, contact the SHRM Foundation at 703.535.6020. Online at shrmfoundation.org

Table of Contents

iii	Foreword
v	Acknowledgments
vii	About the Authors
1	The Business Case for Effective Talent Assessment
2	The Assessment Landscape
2	The Evolution of Assessment Technology
3	Future Trends
4	Tools of the Assessment Trade
4	Job Analysis
5	Assessment Methods
8	How to Select the Right Assessment Method
10	Validity
11	Adverse Impact
11	Tradeoffs Between Validity and Adverse Impact
12	Assessment Utility: Business Benefits
12	Costs
13	Candidate Reactions
14	Legal Issues
17	Conclusion
19	Sources and Suggested Readings
33	Endnotes

FOREWORD



Organizational success is largely driven by the acquisition of the best talent available. However, statistics reveal that 25 percent of mission-critical positions often remain vacant after five months. As organizations compete for the top candidates, they expend large sums of money, time and other resources in advertising and recruitment. Research shows that the careful use of talent assessments can result in increased productivity, cost savings and better retention of top employees.

To help you sort through the many assessments available and find those that will work best for you, the SHRM Foundation has created *Choosing Effective Talent Assessments to Strengthen Your Organization*. This guide focuses on the tools of the trade and how to select the right assessment to achieve your business goals. It helps employers better understand the value of assessment, as well as the legal and technical issues involved. In addition, it offers a clear roadmap to selecting the effective assessment methods to target top talent and boost productivity.

This report is part of the SHRM Foundation's Effective Practice Guidelines series, which now includes more than 20 titles. Created in 2004 for busy HR professionals, the series integrates research findings with expert opinion on how to conduct effective HR practice. It provides the tools to successfully practice evidence-based management. Other recent reports include *Talent Acquisition*, *Creating a More Human Workplace* and *The Aging Workforce*. To ensure the material is both practical and research based, the reports are written by subject-matter experts and are then peer-reviewed by both academics and HR professionals.

The SHRM Foundation champions workforce and workplace transformation and inspires HR professionals to make it happen. Our educational resources are now used in hundreds of college classrooms worldwide. The SHRM Foundation also awards hundreds of scholarships annually to support education, conference attendance and professional certification for HR professionals. All this good work is made possible by the generous support of donors like you.

We encourage you to learn more. Please visit shrmfoundation.org to download other complimentary resources and to find out how you can get involved with the SHRM Foundation. Thank you!

Mark Schmit, Ph.D., SHRM-SCP
Executive Director, SHRM Foundation



ACKNOWLEDGMENTS

The SHRM Foundation is grateful for the assistance of the following individuals in producing this report:

REVIEWERS

Howard J. Klein, Ph.D.

Professor of Management and Human Resources
Fischer College of Business
The Ohio State University

Jim Link, SHRM-SCP

Chief Human Resources Officer
Randstad North America

Kameron M. Carter

Ph.D. Candidate
Department of Management and Organizations
University of Iowa

PROJECT MANAGER

Dorothy Mebane, Ph.D.

Manager, Foundation Programs
SHRM Foundation

The authors would like to thank Emmanuel Rilhac, head of EMEA & Global Functions, CEB, Arlington, Va., and Donna L. Weiss, director, Product Management, CEB, Arlington, Va., and two anonymous reviewers for their comments on an earlier draft of this report. Their comments contributed significantly to shaping the final product.

ABOUT THE AUTHORS



Elaine Pulakos, Ph.D., has spent her career working with organizations to design and implement talent management systems and processes in the areas of staffing and assessment, performance management, leadership development, and succession management. She is a recognized talent management thought leader and frequently invited speaker to address both practitioner audiences and the scientific community. She has authored numerous articles, book chapters and books, as well as three best-practice volumes for the Society for Human Resource Management (SHRM).

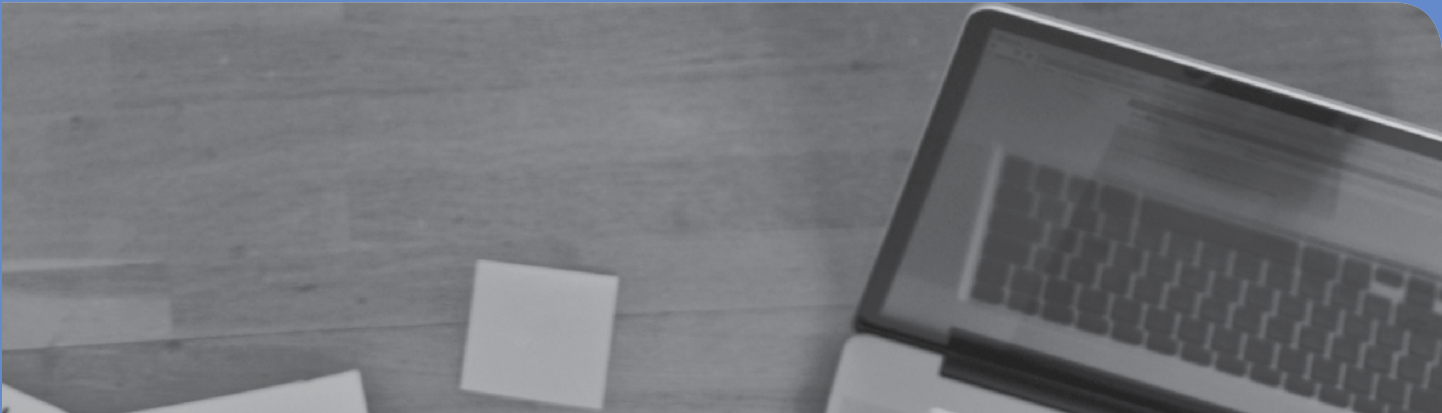
Dr. Pulakos's work has been recognized with several awards, including the Society for Industrial and Organizational Psychology's (SIOP's) Distinguished Professional Contributions Award, the M. Scott Myers Award for Applied Research in the Workplace, and the William A. Owens Scholarly Achievement Award. Elaine is a fellow of the American Psychological Association and SIOP, for which she also served as president. Currently, she leads two groups within CEB: global measurement product development and PDRI, a CEB company.



Tracy Kantrowitz, Ph.D., is research director at CEB Talent Assessment. In this role, she is responsible for the development of assessment content and research related to the use of assessment tools for employee selection and development. Dr. Kantrowitz has published in leading journals and presented at national conferences on topics such as predictors of job performance, computer adaptive testing, unproctored Internet testing and mobile assessment. Tracy's primary areas of expertise relate to test development, validation, research methodology, statistics and psychometrics, job analysis, and selection system design.

Tracy was previously chair of the Professional Practice Committee for the Society for Industrial and Organizational Psychology (SIOP). She has received awards from SIOP, where she was elected fellow, including the M. Scott Myers Award for applied research in the workplace and the Distinguished Early Career Contributions - Practice award.

Dr. Kantrowitz holds a Ph.D. in industrial/organizational psychology from the Georgia Institute of Technology.



Research reveals that careful talent assessment results in significant productivity increases, cost savings and decreased attrition.



CHOOSING EFFECTIVE TALENT ASSESSMENTS TO STRENGTHEN YOUR ORGANIZATION



THE BUSINESS CASE FOR EFFECTIVE TALENT ASSESSMENT

Organizations today compete fiercely for talent. Recent statistics indicate that 25 percent of critical positions remain unfilled after five months.¹ Firms invest an enormous amount of money, time and other resources in advertising and recruiting strategies to attract the best candidates because today's executives understand that one of the most important drivers of organizational success is outstanding talent. Research also reveals that careful talent assessment results in significant productivity increases, cost savings and decreased attrition.

Questions Leaders Are Asking	Organization Challenges
How can we reduce turnover?	On average, one in four new hires leaves within a year. ²
How can I find people who will succeed in specific positions?	Each position has many applicants, but 65 percent of them do not meet basic requirements. ³
How can we identify effective leaders?	Only 12 percent of today's employees are enterprise leaders. ⁴

Given the proven bottom-line value of attracting and retaining talent, why do so many organizations and HR professionals employ ineffective, haphazard approaches to selecting employees? This report addresses how and when to use assessments to improve hiring and retain the most valuable employees in a competitive market.

Although the term "assessment" can include a wide variety of tools, including resume reviews, applications, reference or background checks, and interviews, the focus here is on more formal scientific assessments. After a brief survey of the value of different types of assessments, we will focus on practical explanations of technical and legal issues and then offer a roadmap to choosing the methods that offer the best return on investment for your organization.

THE ASSESSMENT LANDSCAPE

The nature and variety of assessment tools have evolved considerably in recent years as analytics and data have become more important in the business environment. Formal assessments are used to hire new employees, evaluate current employees for different roles, identify employees with high leadership potential, compare talent within an organization against industry or geographical benchmarks, understand talent strengths and gaps, and develop employees' long-term value. Advances in technology mean that more tools are accessible online, on-demand, and with immediate reporting and feedback.

Technology has also given rise to more available data on individuals and new tools that enable organizations to mine these data for competitive advantage. For example, social media and other data collected via technology are used for predictive analytics, modeling trends, passive assessment of candidates and uncovering insights to enhance business performance. The use of immersive, personalized assessment experiences is on the rise.

In the hiring context, organizations are using these new tools and experiences to provide job previews and sell the organization and its brand. Some firms want to provide brand-enhancing hiring experiences that will attract candidates and retain them as customers, even if they are not offered jobs. This trend is important because poor candidate experiences are broadly shared throughout social networks and can yield decreased use of the organization's products and

services.⁵ Approximately two-thirds of HR professionals report that their organizations are putting a premium on positive applicant experiences, although relatively few collect data to monitor reactions to assessment procedures.⁶

One of the most dramatic shifts involves the evolution of the *candidate-driven job market*,⁷ which has changed how organizations compete for talent and the nature of their recruitment and assessment processes. Candidates want brief, informative, and engaging hiring experiences and transparency so they know what to expect, where they stand and how much effort to put into a hiring process. All these new expectations are transforming assessment methods and practices.

The Evolution of Assessment Technology

Unsupervised online testing

has become the standard for assessment administration because it is so efficient, decreasing the time needed to fill vacancies and reducing costs.⁸ Initially, online testing prompted two major concerns: test security and candidate integrity.

Concerns about security arose because early versions of online assessments were static, meaning that all candidates received the same items, which could be copied and shared easily.⁹ If tests are copied and shared, they become invalid and must be replaced, creating unsustainable costs for companies.

Computer adaptive testing (CAT) methods are now the norm because CAT allows for dynamic selection of

Assessment Trends

- Machine learning and analytics are introducing new assessment techniques.
- Globalization requires localized assessments that can easily reach overseas.
- Technology is enabling on-demand, online assessments with immediate results and feedback.
- More firms are using immersive, compelling assessment experiences.
- Organizations want personalized and contextualized assessment.
- Candidate-centric hiring requires shorter, more transparent and engaging candidate experiences.

questions from a large set of items based on how each candidate is responding: If a candidate responds correctly, the next item will be more difficult; if a candidate responds incorrectly, the next item will be easier. A customized assessment with a unique combination of items focuses on each candidate's level of ability and knowledge. CAT creates more secure and more efficient assessments, with fewer items in less time, to meet the demands of candidate-centric hiring.

CAT represents a critical innovation in online assessments¹⁰ but does not resolve the problem of candidate integrity or making sure that the person taking an online assessment is the same person who is applying

Test Equivalence Is a Key Issue

As assessment devices evolve from booklets to PCs to mobile phones and tablets, we need to consider whether the mode of administration, rather than the assessment itself, is leading to lower or higher scores. The ideal situation is that no matter what device is used, the candidate earns the same score.

for a job. To ensure candidate integrity, some organizations ask candidates who have progressed through the hiring process to complete a brief onsite assessment at a convenient point to verify the candidate's online score.

Mobile assessment now provides even more convenience than assessments via desktop computers. Consumer data indicate that mobile devices may be more readily available to underrepresented groups, including African Americans, Hispanic Americans and women.¹¹ Assessment administration on mobile devices may therefore result in larger and more diverse candidate pools.

Differences in mobile devices mean that mobile assessment can be complex, requiring more nuanced evaluation. For example, larger screens on tablets allow some forms of assessment that will not work well on phones. How information is input, whether via touch screen or keyboard and mouse, presents unique challenges. A new class of assessments that use mobile-centric design principles has arisen to optimize the user experience. We are seeing a demand for innovative assessments that go beyond multiple choice and look and feel more like games. The result of these innovations is that assessments are being judged by their design and accessibility

as much as by their measurement and predictive qualities.

Some assessments, including personality and past-experience assessments, show little difference in scores when completed on mobile devices or PCs.¹² But research on cognitive ability assessments is inconsistent, with some studies finding equivalence across device types,¹³ and others finding differences—with lower scores on mobile phones.¹⁴ Overall, however, initial research shows relatively small differences restricted to certain assessment items.

Low-effort assessment uses new technology and machine learning to sift through large amounts of unstructured information, including resumes and applications, to collect predictive information about candidates. The information is coded and validated against work outcomes, such as job performance and retention, to produce algorithms that managers can use to inform their hiring decisions. These techniques automate identification and scoring of predictive information so that large numbers of candidates can be screened in minimal time at low cost.

Low-effort assessment is still in its early stages, so all the advantages and disadvantages are not yet clear, but it is probably best suited for situations in which past experience is especially

Technological Advances in Assessment

- Unsupervised online testing.
- Computer adaptive testing.
- Mobile assessment.
- Low effort assessment.

important in predicting future performance.

Future Trends

Technology will continue to spark revolutionary changes in assessment methods and practices. For example, the ability to passively target candidates from available information sources could lead to transformational changes in hiring processes. Today, applicants apply for positions based on job postings, but tomorrow organizations may search data for key talent in ways that bypass postings. This is uncharted territory and will require new policies on privacy, legal and ethical considerations. As hiring practices span global markets, human resource professionals will also have to become familiar with additional laws and regulations.

Perhaps most important, early adopters of new technology should be aware that without sufficient research behind it, an innovation might undermine rather than help an organization to make effective, defensible hiring decisions. In some countries, litigation and audits of hiring practices will lead to new standards and precedents related to new assessment technology, and organizations must be alert for these changes. Later in this

report, we will look at the science behind assessments, legal requirements and workers' rights to help organizations navigate these choppy new waters.

TOOLS OF THE ASSESSMENT TRADE

Organizations can use a broad array of techniques to make selecting the right assessment tool easier. The first order of business is to do a thorough job analysis, followed by a review of the various types of assessment methods available.

Job Analysis

The first step in selecting an assessment tool is to understand what the target job or role requires the employee to do by conducting a *job analysis*¹⁵ that focuses on identifying *critical work activities* and/or *competencies*.

An in-depth discussion of job analyses is beyond the scope of this report, but they typically

Investigator Job: Sample Work Activities

- Provide testimony by stating facts and answering questions.
- Gather and review information to obtain evidence, or develop background information on subjects.
- Integrate information to uncover relationships between individuals, events or evidence.
- Work in a team environment as a team member or leader.
- Calm and reassure victims and others in tense situations.

Investigator Job: Sample Competencies

- Communication: Speak and write clearly and concisely, with self-confidence and using effective inflection and gestures.
- Critical thinking: Question assumptions and identify merits in logic.
- Collaboration: Elicit cooperation from other individuals and organizations.
- Professionalism: Maintain a professional demeanor and appearance at all times.
- Handling stress: Remain calm and levelheaded.

Job Analysis Steps

STEP 1. Observe or interview job experts to develop a list of work activities and competencies workers need.

STEP 2. Ask job experts to rate which work activities and competencies are most critical for effective job performance.

STEP 3. Analyze the data to prioritize the most critical work activities and competencies for your organization.

STEP 4. Select assessments that measure the performance of the most critical work activities or competencies required on the job.

include four major steps, as outlined above.

Every organization's unique character is reflected in how it performs a job analysis and then chooses assessments. Business strategy, organizational culture and contextual factors always play major roles. Contextual factors may include high-versus-low growth, high-versus-low change and high-versus-low team interdependency. A recent study of thousands of organizational leaders across 85 companies showed the importance of organizational context factors in successful identification of leaders at all levels.¹⁶ These results reinforce the value of thinking beyond job-specific competencies

to broader performance requirements.

One useful source of job analysis information is O*NET, a database that is continually updated by surveying a broad range of workers from hundreds of occupations. Employers and job seekers can leverage the O*NET database to obtain information for virtually any job in the global economy. The database forms the heart of O*NET OnLine, an interactive application for exploring and searching occupational information (<http://onetonline.org>). Of course, it is always best to verify that job information from O*NET fits the specific job and situation in your organization.



Assessment Methods

After you understand job requirements, you can select assessments that measure the most critical work activities or competencies. For example, is the assessment goal to select someone to perform a specific job in the near term or to predict employee potential for future leadership roles? In the former case, the assessment can be targeted to the most critical competencies for the job. In the latter case, a broader and more comprehensive whole-person assessment will probably be more useful.

Assessment types run along a continuum from more observable, direct measures of performance to less observable, indirect measures of performance. An iceberg analogy helps illustrate the different assessment types.

Work behavior is the tip of the iceberg—the observable performance on the job. Work behavior assessments require candidates to perform simulated aspects of the job and provide an observable measure of performance that should forecast real on-the-job performance. These assessments should correspond closely to the specific work that candidates will perform and tend to be expensive, time-consuming and somewhat limited in their use.

Acquired knowledge, skills and experience sit just below the surface. Although less directly observable than work behaviors, assessments of job-relevant knowledge, skills and experience are extremely effective predictors of job performance. Because these assessments require measuring unique capabilities relevant to *specific* jobs, they are not generally applicable *across* jobs.

Innate ability and attributes, including personality and interests, lie deep below the surface and also have been proven to be effective predictors of job performance. Because assessments of innate characteristics do not focus on job-specific traits, they are the most flexible and applicable for use across different jobs.

Work Behavior Assessments

- Behavioral interviews.
- Work simulations.
- Physical ability tests.

Acquired Knowledge, Skills and Experience Assessments

- Bio data/experience.
- Job knowledge.
- Writing skills.
- Situational judgment.

Innate Ability and Attribute Assessments

- Cognitive ability.
- Personality.
- Integrity.
- Physical fitness.

Below we discuss *specific assessment methods*, all of which are backed by strong research over decades. They are all readily available and effective for small and large companies. We do not include some newer assessment methods with less research available to back them up.

Cognitive ability assessments typically consist of multiple-choice items that measure various mental abilities, such as verbal and numerical ability, reasoning, and reading comprehension. These assessments have proven to be extremely useful predictors of job performance and are used frequently in hiring for many types of jobs.¹⁷ Some measure distinct abilities, whereas others measure a

Example of Cognitive Ability Test Items		
Verbal ability	Numerical ability	Reasoning ability
Innocuous means the same as: a. Harmless b. Preventative c. Distasteful d. Futile	16% of 62.5 is: a. 8.44 b. 8.44 c. 0.84 d. 8.4	1 3 2 4 3 5 4 6 5 a. 4 b. 5 c. 6 d. 7

combination to provide a measure of general mental ability or “g.”¹⁸

Cognitive ability tests administered online in unsupervised testing situations often employ adaptive methods. The mobile environment poses some unique challenges because the typical questions are too complex and lengthy for mobile phones, although they may work well on tablets. Newer test designs will probably be optimized for mobile administration.

Personality assessments

measure stable, job-relevant characteristics that have been shown to be effective predictors

Examples of Personality Test Items

Standard format:

I have remained calm in situations in which others have become upset.

- Definitely true
- Somewhat true
- Neither true nor false
- Somewhat false
- Definitely false

Format that Mitigates Intentional Desirable Responses

Choose which item is most like you and which is least like you.

- I persist with tasks until completed.
- I treat customer complaints with highest priority.
- I deal well with extremely stressful situations.

of job performance,¹⁸ particularly in combination with measures of cognitive ability. In work situations, conscientiousness, extraversion, agreeableness, openness to experience and emotional stability are measured most frequently.²⁰ Research suggests that conscientiousness is the most robust and useful predictor of performance across jobs. Achievement orientation is also a strong predictor in many jobs.²¹

Personality assessments typically consist of several multiple-choice or true/false items that measure a variety of personality factors. Desirable answers are often quite obvious in this format, so there is considerable concern about how easy it is to “fake” a positive impression. On the other hand, even when candidates do respond with what they assume are the desired answers, personality assessments are still useful for predicting job performance.²² Some assessments avoid the problem by using a format that presents equally desirable response options representing different personality traits. This item format—asking candidates to choose which answer is *most* like them—is less transparent than a simpler scale. These assessments work well across different devices.²³

Integrity assessments measure attitudes and experiences related to honesty, trustworthiness and dependability.²⁴ They are typically presented in a multiple-choice format. Similar to personality assessments, integrity assessments can be easily administered across different devices without problems.

Physical fitness tests measure overall fitness, strength, endurance or other capabilities necessary to

Example of an Integrity Assessment Item

It is okay to misrepresent the truth if being completely honest will create problems that small “white lies” can solve.

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree

Examples of Physical Fitness Tests

- Run a mile.
- Lift two 20-lb. weights for 50 repetitions.

perform a job. By nature, they don't work in an online environment.

Biographical data (bio data)

assessments ask candidates questions about their past experience in areas that have been shown to predict future performance.²⁵ For example, a question about the number of volunteer organizations to which someone has belonged could indicate the extent to which he or she is willing to pitch in and help others. A question about where one's spending money came from in high school could indicate a history of responsibility and independence.

Some bio data assessments are scored using data-driven techniques, meaning the points associated with each response option are determined according to the relationship between that response and an important job outcome. Other bio data

assessments are rationally scored, meaning the test items are scored according to the degree to which each response option measures an important competency for a particular job. Biographical inventories consist of multiple-choice items, and like personality assessments, they tend to be equivalent across different modes of administration.²⁶

Another form of bio data is an assessment called an *accomplishment record*. In this case, candidates prepare a written account of their best accomplishments in key competency areas, such as planning and organizing, customer service, and conflict resolution.²⁷ Candidates also provide the name of an individual, such as a past supervisor, who can verify each accomplishment. Evaluators are trained to score the accomplishments in a consistent manner using standard rating criteria. Sometimes

Examples of Bio Data Assessment Items

To how many volunteer organizations do you belong?

- a. 0
- b. 1
- c. 2 to 4
- d. 5 or more

Where did you get your spending money in high school?

- Allowance from family
- Own earnings
- Partly allowance, partly earnings
- Other sources
- Had no spending money

automated scoring algorithms are used. Accomplishment records are effective predictors of job performance.²⁸

Job knowledge assessments

measure critical knowledge areas—usually technical knowledge—needed to perform a job effectively.²⁹ These assessments are appropriate to use when candidates must possess a specific body of knowledge prior to job entry, not when candidates will be trained on the job. Job knowledge tests typically consist of multiple-choice items, but they sometimes include essay items. In addition to hiring assessments, job knowledge tests are used often for certification and licensing assessment.

Knowledge tests can be administered well online using large banks of items and adaptive technology, but there is no definitive research about whether they can be administered effectively on mobile devices. Certification and licensing assessments, especially those that are high stakes, are still administered in a supervised testing environment for the most part.

Automated writing assessments

can be developed to measure various competencies beyond the writing itself, including a candidate's persuasiveness, critical thinking, attention to detail, commercial focus and political savvy. Technology is used in lieu of human scoring with a technique called *latent semantic analysis*. Writing assessments are especially useful in selecting people to fill positions with significant writing demands, such as attorney, analyst or investigator.

Development of the writing prompts and scoring algorithms requires quite a bit of time and resources, but the efficiency and cost savings

Example of a Job Knowledge Item

The principle of the “lever” is essential to the use of a:

- a. Hydraulic jack
- b. Plow
- c. Car steering wheel
- d. Forklift

How Are Writing Assessments Created?

First, administrators prepare a written prompt and give it to about 200 candidates. At least two human evaluators independently score each writing sample using defined ratings scales and come to an agreement. These scores and scales are used to train automated scoring algorithms, which are then used to score later writing samples.

after the initial investment can be significant. This type of assessment is suitable for online environments as long as it is possible to confirm that the candidate, and no one else, is the one providing the writing sample.

Situational judgment tests present candidates with situations just like those they would encounter on the job, with viable options for handling each situation.³⁰ Candidates might be asked to select the most effective, or the *most and least* effective, ways of handling a situation from a few options. These tests are more expensive and complicated to develop than many other types because developing scenarios with several equally attractive,

Example Situational Judgment Item

You have just prepared a report and checked it for accuracy. Before you attend a meeting to submit your report, you review the typed version and note many serious errors. In this case, you would:

- a. Show the original and the typed version to the person in charge of typing and demand that the errors be changed before the meeting.
- b. Present the report at the meeting, point out the errors and explain the typist made them.
- c. Present the errors to the typist, ask him or her to make the corrections and explain to those at the meeting that the report is still being typed.
- d. Present your report at the meeting without mentioning the errors. Notify everyone of the corrections after the meeting.

viable responses is difficult. They are typically administered in written or videotaped form, and both formats can be administered online, but administering them on mobile devices with smaller screens may be challenging.

Structured behavioral interviews—unlike more traditional, unstructured interviews—consist of specific questions that assess critical competencies required for specific jobs.

Interviews are among the most common hiring methods used in all sorts of organizations.³¹ Unfortunately, there are rarely any agreed-on standards for evaluating responses, and unstructured interviews are not particularly useful for predicting job performance.

An important characteristic of any effective interview is that it provides *standardized rating criteria* to help interviewers judge the quality and effectiveness of candidate responses. Interviewers must learn to conduct an interview properly, probe for additional information and apply the rating criteria accurately.

Structured behavioral interviews can assess almost any aspect of competency, but employers most

often use them to assess qualities like teamwork, communication, leadership, planning and adaptability.

Work simulations measure how a candidate performs activities that the job requires, and of course, employers should use them only when knowing how to perform those activities is truly a prerequisite.³² These simulations can be time-consuming and expensive to develop. Sometimes, trained evaluators do the scoring, but today assessment and feedback are often automated.

Physical ability tests are used to select people for physically demanding jobs, such as police officer, firefighter or warehouse worker.³³ These assessments require candidates to perform actual work tasks to determine whether they can meet the physical requirements of a job. The tests are often scored on a pass/fail basis and are often timed. Physical *ability* tests are specific to work activities, whereas physical *fitness* tests measure more general overall physical health, strength and endurance.

Structured Behavioral Interview Question and Sample Rating Criteria

Tell me about a time when you were able to establish rapport with someone in a situation that made it difficult to do so. What were the circumstances? What did you do? What were the results?

Low 1	Moderate 3	High 5
Made minimal attempts to understand the person's perspective. Developed only a surface-level relationship in a simple situation.	Attempted to understand the other person's perspective. Developed a positive working relationship with the person in a moderately difficult situation.	Effectively reached out and sought to understand the person's views. Developed an extremely positive relationship with the person in an extremely difficult situation.

Sample Physical Ability Tests

Firefighter

- Climb a ladder while carrying equipment.
- Drag a hose.
- Carry a person down from a building.

Warehouse Worker

- Lift materials of weight that must be lifted on the job.
- Drag materials for distances required on the job.
- Roll barrels.
- Carry materials the distance they must be carried on the job.

HOW TO SELECT THE RIGHT ASSESSMENT METHOD

In this section, we provide guidelines to help human resource professionals make informed decisions about assessments. We will look at four important criteria, and the answers will be based on solid research over time:³⁴

- **Validity:** Is the assessment method useful for predicting job performance?
- **Adverse impact:** Do members of protected groups, including minorities, women and employees over age 40, systematically score lower than others?
- **Cost:** What is the expense of developing, administering and maintaining the assessment?

- **Candidate reactions:** Do candidates react positively or negatively to the assessment?

The results, explained in the table below, demonstrate that no simple formula for choosing the right assessment exists because each type measures specific aspects, and each has advantages and disadvantages. Often, a combination of assessments works well to optimize results based on your organization’s specific goals and needs. Assessments can be phased in—using a shorter one first, to screen out the most unqualified candidates, and then using more complex methods for higher-potential candidates.

Evaluation of Assessment Methods on Four Key Criteria				
Assessment Method	Validity	Adverse Impact	Costs to Develop/ Administer/ Maintain	Candidate Reactions
Cognitive ability	High	High (minorities)	Low-high/low/high	Somewhat favorable
Personality	Low-moderate	Low	Low-high/low/low	More favorable if job-relevant
Integrity	Moderate-high	Low	Low/low/low	Less favorable
Physical fitness	Moderate to high	High (females and over age 40)	High/high/low	More favorable
Biographical data	Moderate	Low-high for different types	High/low/moderate	More favorable if job-relevant
Job knowledge	High	High (minorities)	Low-high/low/high	More favorable
Writing	High	High (minorities)	High/low	Somewhat favorable
Situational judgment	Moderate	Moderate (minorities)	High/low/moderate	More favorable
Behavioral interviews	High	Low	High/high/low	More favorable
Work simulations	High	Low-high, depends on what is measured	High/high/high	More favorable
Assessment centers	Moderate to high	Low-high, depends on what is measured	High/high/high	More favorable
Physical ability	Moderate to high	High (females and those over age 40)	High/high/low	More favorable

What is an assessment center, and when do you need one?

An assessment center offers a specific type of work simulation, usually related to executive, managerial, or supervisory jobs. Candidates respond to situations, problems, and tasks in role-play exercises, “in-box exercises” that ask them to prioritize and make decisions based on incoming information, case exercises, and group discussion. Video and voice recordings mirror what candidates see and hear on the job. More sophisticated options, such as branching role-plays, can automate personal communication by providing adaptive responses based on what the candidate does or says.

In some cases, assessment centers can also be used for non-managerial job candidates. “Day in the life” simulations ask job seekers to handle scenarios in workflows similar to what happens on the job. These simulations provide robust information for those making hiring decisions for the candidates. As they become easier, cheaper, and more flexible, use of simulations may increase over time, especially in latter stages of a hiring process when people tend to be more willing to invest in a longer, in-depth commitment.

Validity: The Most Important Consideration

Does the assessment predict how individuals will perform on the job?

Criterion Validity

- Determines the strength of the relationship between scores on an assessment and job performance effectiveness, tenure or other important outcomes.
- Validity is measured using a correlation coefficient, which ranges from 0.0 to 1.00.
- Assessment validities typically range from .10 to .40.
- A minimum validity to target is .25 or above.

Content Validity

- Involves providing evidence that an assessment is a direct measure of work behaviors.
- Based on expert judgment.

You may find the array of possibilities requires you to seek advice from an assessment expert. Below we explain some of the critical information to keep in mind.

Validity

Validity means that an assessment is job-relevant and useful for predicting performance. Clearly, this is the most important consideration in selecting an assessment. The two most common types of validity human resource professionals encounter are *criterion* and *content* validity.³⁵

Criterion validity involves demonstrating a useful relationship

between scores on an assessment and important job outcomes. For example, assume that a sales job requires exceptional customer service. If the assessment is a *valid* predictor of job performance, candidates who score higher on customer service will perform better, with higher sales and higher customer loyalty ratings, and those who score lower will not perform as well. To establish criterion validity, a firm would evaluate the strength of the relationship between performance on the assessment and job performance on a scale of 0 to 1.00—although scores would never reach a perfect 1.00 because so many other factors are at play.

Tip: Look at Validation Studies

In considering the use of commercially available assessments, HR professionals should look for how many criterion validity studies have been conducted on the assessments and meta-analysis results, if available. The more successful validation studies a vendor can produce from other organizations and the larger the number of individuals included in validity studies, the more confidence you can have in the quality of the assessment.

Example of Criterion Validity Studies

- Administer the assessment to be validated to a large group of employees (at least 300).
- Collect measures of these same employees' performance on the job.
- Correlate the assessment scores with the job performance scores, using a correlation coefficient.
- The correlation is the strength of the relationship between the assessment and job performance.

After multiple validation studies have been conducted for a given assessment type, summarize the results across studies, and make a stronger statement about the true validity of the assessment. In the table above, we analyzed validation studies and used them to identify "low" (.20 or less), "moderate" (.20 to .40) and "high" (.40 to .50) validity.

Assessments with low-to-moderate validities can add value, especially when they are used in combination with other assessments. For example, cognitive ability and personality are largely unrelated, but, cumulatively, they combine to produce higher validity than when used alone.

Content validity involves demonstrating that an assessment provides a direct measure of critical work activities or behavior. Comprehensive job analysis information and expert judgment are used to develop work simulations, ensuring that the assessment closely mirrors job requirements.

Comparisons of criterion and content validity reveal that criterion validity is stronger and tends to be viewed as more credible. However, content validation is perfectly acceptable for demonstrating the job relevance of work samples, simulations and other direct measures of work behavior. Criterion validity can be used to validate any type of assessment and is the only type of validity that can be used when an inference needs to be made about the relevance of an assessment for predicting job performance instead of using a direct measure of performance. Most assessments, including cognitive ability, personality, bio data and integrity, require inferences about their relevance for predicting performance, so criterion validation is the most common form of validity.

Successful criterion validation relies on having performance or other such outcome measures that accurately distinguish employees' effectiveness on the job. Performance ratings that are available in human resource information systems (HRIS) or performance management systems are often of such poor

Looking at the Numbers: Effect Size and Adverse Impact

An *effect size* can be calculated to compare the assessment scores between two groups. Effect sizes range from 0, indicating no difference between groups, to 1.00 or more, indicating a large difference. Effect sizes less than .20 are considered small, and those near .80 or more are considered large, signaling a likely adverse effect. Understanding effect sizes is useful because they are often reported in test manuals and research reports evaluating assessments.

quality that they do not enable successful validation.

The bottom line is that criterion validity studies are so complex that conferring with an assessment expert is best if you need to do criterion validation work.

Adverse Impact

Adverse impact is a concept defined by the U.S. government to establish evidence of discrimination. It occurs when the results of an assessment or hiring lead to a disproportionately

Adverse Impact

An assessment has adverse impact if it results in hiring proportionately fewer protected group members than majority group members.

Calculating Adverse Impact

The 80 Percent Rule:

- 50 females are assessed and 30 females are hired, or 60 percent.
- 50 males are assessed and 45 males are hired, or 90 percent.
- The ratio of females to males hired is 67 percent.
- A ratio of < 80 percent is evidence of adverse impact.

small number of protected group members (e.g., racial minorities, females, individuals over age 40) being selected. If the proportion of protected group members hired is less than 80 percent of the proportion of majority group members hired—referred to as the *80 percent rule*—adverse impact is suspected. If an assessment produces an adverse impact, the only way to defend its use is by showing that it is a valid predictor of performance and that alternative measures are not feasible.

The 80 percent rule is not the only way adverse impact is defined or

Strategies to Reduce Adverse Impact

What To Do

- Recruit more qualified candidates from protected groups.
- Assess a wider set of competencies, such as contextual performance (e.g., citizenship, helping others, expending extra effort), which is predicted by measures that produce minimal adverse impact (e.g., personality and bio data).
- Use assessments with more adverse impact later in the process, although this approach is likely to help only when there are many applicants per position.

What Not To Do

- Administer test preparation programs. They may increase candidate comfort but do not reduce adverse impact.
- Remove test items with large group differences.

evaluated. The U.S. Office of Federal Contract Compliance Programs uses an alternative approach, one to which commercial businesses of a certain size, with federal contracts, must comply. With the exception of South Africa, the concept of adverse impact is not widely used outside the United States, but other countries have specific laws, regulations and compliance requirements that govern hiring practices and employee rights.

Adverse impact does not only happen in the assessment process. Resume screening, reference checks, interviews and recruitment practices can also result in an adverse impact against protected groups. HR teams should not shy away from using assessments just to avoid adverse impact, especially given the improved performance and business outcomes that proven assessments provide.

Tradeoffs Between Validity and Adverse Impact

All organizations want the best-performing workforce, and most also want a diverse workforce that does not exclude protected groups.

While evidence of validity can be used to justify and defend assessments that produce an adverse impact, most organizations want to mitigate adverse impact as much as possible. Some assessments that produce high levels of validity also produce high adverse impact. A number of strategies may reduce adverse impact without reducing validity, including enhanced recruiting of protected groups, assessing a broader array of competencies beyond the technical to include motivational and behavioral factors,³⁶ and adjusting the placement of assessments within the hiring process.³⁷ Less effective

Assessment Utility

What are the business outcome benefits of using assessments?

Selection Success Gains

- Number or proportion of successful candidates before and after implementing a valid assessment

Business Outcome Gains

- The financial consequences of using a valid assessment
- Revenue and profitability increases compared to assessment costs

strategies include providing test preparation sessions for candidates³⁸ and removing test items that have larger group differences.³⁹ While some strategies have shown promise, elimination of adverse impact entirely is still a challenge and is something that organizations should monitor and manage across all talent management processes, systems and decisions.

Assessment Utility: Business Benefits

Assessment utility addresses the business benefits of using assessments and is helpful in explaining the value of assessment to decision-makers who are outside the human resource team. Two types of utility are apparent in assessments: selection success gains and business outcome gains.⁴⁰

Selection success gains compare successful candidates before and after assessment implementation. The greater the selection success gain, the higher the assessment utility.

Business outcome gains are more interesting, particularly the revenue and margin gains that result from selecting higher-performing employees compared with the assessment costs. The formulas used to evaluate business outcome gains are more complicated than simply comparing revenue/profits versus costs. They consider additional factors, including how effectively the assessment predicts performance, the number of applicants and tenure of the selected group.⁴¹

When decision-makers see the positive impacts of assessment in terms of key business outcome metrics, the value of assessment usually becomes more apparent and compelling.

Costs

Some assessment methods involve higher costs to develop, maintain and administer than others, although technology has significantly reduced the large variability in costs. Mobile administration should further reduce and equalize administration costs. In terms of development, automated work simulations tend to be the most costly to develop because they typically require test experts to work with job experts and IT developers on design and automation. Work samples that are not automated can be even more expensive than automated assessments because they involve equipment, facilities and training evaluators to observe and score candidates. At the other extreme, online multiple-choice assessments of innate characteristics can usually be developed independently by testing experts and are highly efficient to administer and score.

The buy-versus-build decision is key for any organization. Should you license standard assessment content or develop and host custom

Buy Versus Build?

Buying Commercial Assessments

Pros:

- Implementation is fast.
- Assessment quality has been demonstrated.
- Vendors provide hosting, maintenance and updating.
- Benchmarks provide talent insights.
- Localized assessments and norms may be available that can be used globally.

Cons:

- Licensing agreements yield ongoing expense that may be prohibitive.
- Use across situations may raise security or competitive advantage concerns.

Building Custom Assessments

Pros:

- Full customization to the organization's jobs and strategy is possible.
- They present maximum relevance and potential competitive advantage.
- They have no ongoing licensing fees.

Cons:

- Ongoing costs exist to host, upgrade and maintain.
- Significant effort and data are needed to localize assessments and develop appropriate norms.
- Organizations need specialized expertise to build, host and maintain assessments.
- They are time-consuming.

assessments in-house? There are issues of time, cost and convenience to consider in each case.

Some organizations prefer *custom assessment* content that is uniquely designed and validated for their specific situation and embeds the organization's values, culture and context. Using a custom assessment is most important when the assessment is intended to serve as part of a larger, highly branded hiring experience or as both a job preview and an assessment. Some organizations believe custom assessment provides a competitive advantage and better identifies candidates

who are best matched to the organization and job.

After development, custom assessments need to be administered and maintained, which requires ongoing investments in online hosting, maintaining databases of test scores, analyzing and reporting data for decision-making, revising items, and potentially revalidating in the future. If an organization's global assessments have to be adapted for use in different countries, the firm should ensure that the content is adjusted and nuanced so candidates in different cultures can understand it. Effectively managing a custom assessment program

requires specialized expertise among in-house staff or external expert consultants.

Organizations may choose to use *commercially available assessments*, which assessment providers can implement quickly and host, update and maintain. Commercial assessments should always come with evidence of assessment quality across organizations. Some commercial assessment providers offer benchmarking data against which organizations can evaluate how their talent stacks up in their industry or geographical area. Some providers have also amassed large volumes of data globally, which allows them to provide localized assessments with appropriate norms. Some providers can also customize products to meet a specific organization's needs.

In setting up assessment programs, employers should seek the guidance of assessment experts. In many organizations, assessment programs are executed through a combination of in-house staff and external experts or vendors. SHRM provides comprehensive and valuable information about available assessments and related topics through the SHRM Testing Center (shrm.org/testing/).

Candidate Reactions

As many organizations move toward candidate-centric hiring, candidates' reactions have become increasingly important in judging assessments. Candidates who have a positive hiring experience report higher levels of satisfaction and are more likely to recommend a company to others.⁴²

The most critical factors influencing applicants' reactions boil down to whether they believe the organization provided:

- An overall great candidate experience.
- Helpful information and feedback.
- Assessments relevant to the job.

In general, candidates favor assessments that measure work behavior directly through methods like work samples and simulations rather than multiple-choice tests of abstract concepts. Cognitive assessments tend to be viewed more favorably than personality, integrity and bio data assessments.⁴³ Job candidates view assessment items aimed at diagnosing personality or clinical disorders much less favorably than items that measure job-relevant personality characteristics like conscientiousness and agreeableness.

Candidates want information about the job, the organization and how they are doing in the hiring process, including how they could have performed better—but this is a tricky matter. Employers should provide as much information as is reasonable, but discussing assessment performance in detail can divulge information that could lead to security breaches. And providing in-depth feedback for all candidates would be extremely expensive.

A final topic relevant to candidate reactions concerns views of mobile assessment. Some studies have found no differences in perceived fairness moving from PCs to mobile devices,⁴⁴ but others have come to the opposite conclusion.⁴⁵ Research has shown that satisfaction increases as screen size increases, irrespective of the content. Item

Candidates React Positively When:

- They are able to demonstrate job-relevant capabilities.
- Assessment procedures are administered consistently and fairly.
- Feedback is provided so candidates can determine their standing and decide whether to proceed.
- They are given insight into the nature of the job and organization.

types with excessive words were viewed less favorably on mobile devices compared with graphic and image-based items. As mobile technology expands and assessments are designed to take advantage of more mobile features, such as swiping instead of clicking, additional research will be needed to fully understand candidate reactions.

LEGAL ISSUES

Human resource professionals who administer assessments should have a basic understanding of the legal issues and requirements that govern their use. In the United States, two pieces of legislation are especially helpful: Uniform Guidelines on Employee Selection Procedures and the Americans with Disabilities Act of 1990 (ADA).

The Uniform Guidelines specify requirements for selection systems covered under the Civil Rights Acts and under Executive Order 11246 (see <https://www.eeoc.gov/laws/regulations/>

The Importance of Compliance

Some managers and others within your organization may believe that using informal selection methods instead of formal assessments will avoid legal issues and eliminate the need to comply with federal regulations. This is not the case. If any aspect of the hiring process clearly produces an adverse impact, an organization will be required to demonstrate validity if challenged or audited. In practice, well-developed and validated formal assessments actually help mitigate legal risk and exposure.

If members of your firm view legal requirements as burdensome, let them know that the procedures in the Uniform Guidelines are, in fact, the same procedures an assessment expert would employ if the only goal were to identify the best-qualified candidates for a job. Thus, compliance with legal requirements in this case is synonymous with best practices and will yield the best outcomes for any organization.

for the full text of the Uniform Guidelines). These requirements address the question of whether a procedure is causing an adverse impact, and if so, which validation requirements should be met before using the procedure. Two documents can assist in applying the Uniform Guidelines:

- The Society for Industrial and Organizational Psychology's *Principles for the Validation and Use of Personnel Selection Procedures* (see <http://siop.org>).
- *Standards for Educational and Psychological Testing*, developed jointly by the American Educational Research Association, the American Psychological Association and the National Council on Measurement in Education (see <http://apa.org/science/programs/testing/standards.aspx>).

The Americans with Disabilities

Act created substantial requirements for providing reasonable accommodations to applicants with disabilities. If a job requires a skill that a person with a disability is unable to perform, an organization is not required to provide reasonable

accommodation in assessing that skill. However, if there are ways to accommodate disabilities both on the job and in the assessment process, an organization must do so. Examples of reasonable accommodation include:

- Substituting an oral assessment for a written one.
- Providing extra time to complete an assessment.
- Offering an assessment in Braille.
- Providing items in larger print.
- Providing an interpreter who can communicate items in sign language.

Organizations increasingly are held accountable for ensuring that recruitment and selection processes are fair and accessible, and research is slowly accumulating on the fairness of selection tools and accommodations for people with disabilities.⁴⁶ For example, research on longer exam times found that they generally improved the performance of all students, not just those with disabilities.

Several sources offer guidance to help HR professionals

Federal Contractors and Subcontractors: Pay Attention to Recent Regulatory Guidelines for Hires with Disabilities

Regulatory guidelines updated in 2013 reflect increased emphasis on hiring individuals with disabilities, and hiring goals are now in place for federal contractors and subcontractors, with stipulations for actions to be taken in the areas of recruitment, training, record-keeping and policy dissemination. The guidelines call for voluntary disclosure from job candidates to self-identify as individuals with disabilities or protected veterans. The guidelines also require workforce surveys to gather this self-reported information. The data are used to create action-oriented programs.

understand and meet their responsibilities under the ADA. One of the most comprehensive sources of information is the *Job Accommodation Network* (see <http://askjan.org>), maintained by the U.S. Department of Labor, which provides free, expert and confidential guidance on workplace accommodations and disability employment issues.

Validity means that an assessment is job-relevant and useful for predicting performance. Clearly, this is the most important consideration in selecting an assessment.

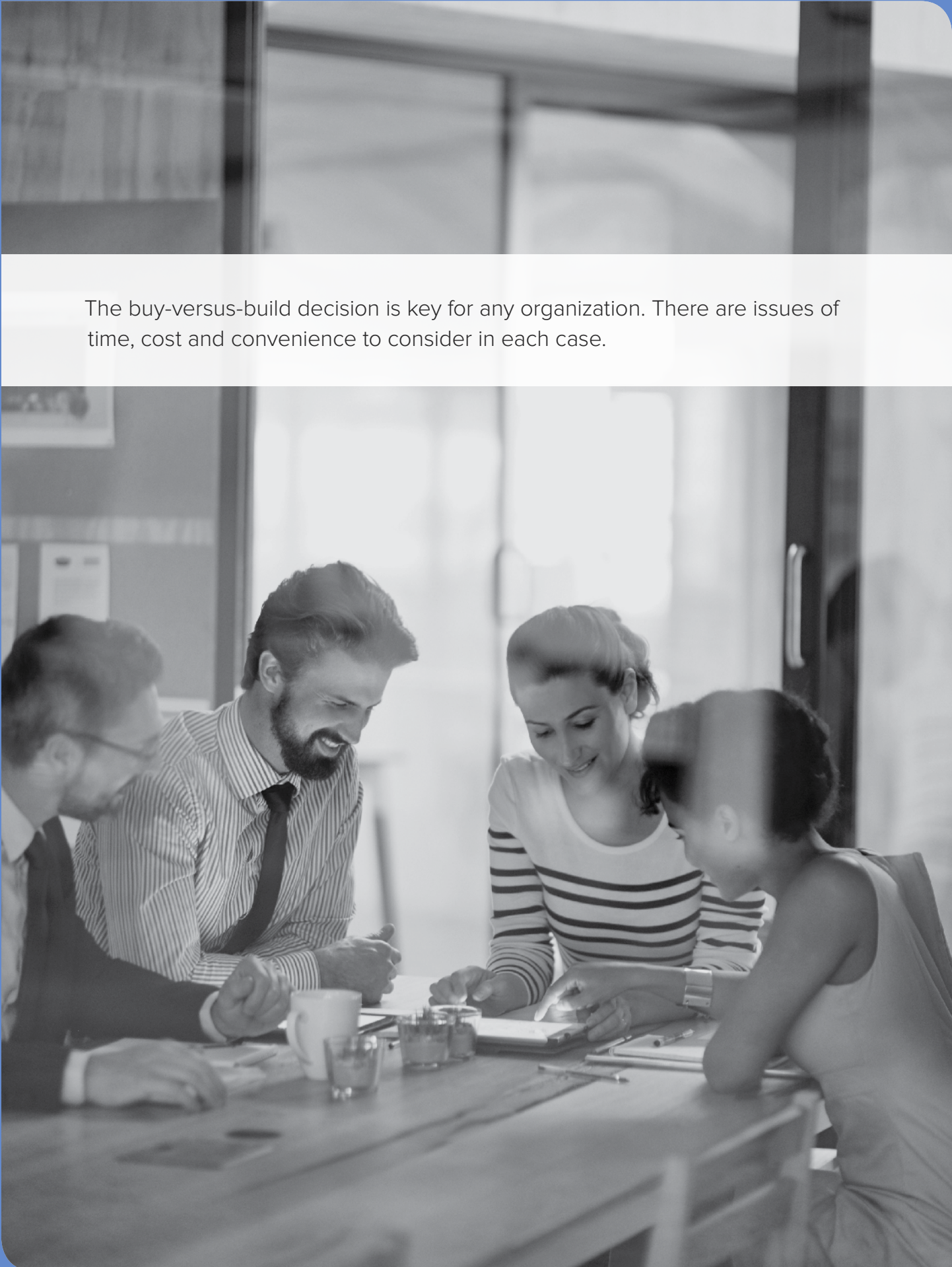




CONCLUSION

In recent years, technological advances, competitive demands for talent and candidate-centric hiring have created powerful commercial forces that are driving changes in assessment methods. Many organizations are demanding shorter, more engaging candidate experiences and less obtrusive assessment methods to help them target talent. Innovative assessment techniques are also being developed. These factors, coupled with continued technological advancement, will almost certainly lead to revolutionary and rapid changes in the assessment field.

Always remember, however, that the primary criterion for evaluating assessments does not change. Managers and human resource professionals must continue to seek effective assessments that will result in talented candidates with strong performance on the job—employees who bring long-term value to the organization.



The buy-versus-build decision is key for any organization. There are issues of time, cost and convenience to consider in each case.

SOURCES AND SUGGESTED READINGS

JOB ANALYSIS

Brannick, M. T., & Levine, E. L. (2002). *Job analysis*. Thousand Oaks, CA: Sage.

This is a user-friendly book about how to conduct a job analysis and use the results. Topics covered include the building blocks of job analysis methods, work (task)-oriented methods; worker knowledge, skills and abilities (KSA) methods; hybrid methods that combine both task and KSA job analysis; job analysis for management jobs and teams; job analysis and the law; doing a job analysis study; and uses of job analysis in job descriptions, performance appraisals, job evaluation, job design, staffing and training.

Gael, S. (Ed.). (1988). *The job analysis handbook for business, industry, and government*. New York, NY: Wiley.

The author argues that job analysis is the hub of virtually all human resource administration and management activity and necessary for the successful functioning of organizations. The handbook addresses job analysis topics and issues comprehensively, with material relevant to a wide audience of practitioners working in business, industry, labor unions, universities, the military, and federal, state and local governments.

UNPROCTORED AND MOBILE ASSESSMENT

Beaty, J. C., Dawson, C. R., Fallaw, S. S., & Kantrowitz, T. M. (2009). Recovering the scientist-practitioner model: How IOs should respond to UIT. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 2, 58-63.

This commentary raises organizational examples and mitigating factors associated with unproctored Internet testing (UIT) in three areas: cheating, ethics and legal concerns. The authors encourage I/O practitioners to execute a research agenda associated with UIT.

Beaty, J. C., Nye, C., Borneman, M., Kantrowitz, T. M., Drasgow, F., & Grauer, E. (2011). Proctored versus unproctored Internet tests: Are unproctored tests as predictive of job performance? *International Journal of Selection and Assessment, 19*, 1-10.

This paper examines the criterion-related validity of unproctored assessments. Results generally indicate that assessments administered in proctored and unproctored settings have similar validities.

Gutierrez, S. L., & Meyer, J. M. (2014). The mobile revolution: Measurement equivalence and mobile device administration. In T. Kantrowitz & C. M. Reddock, *Shaping the future of mobile assessment: Research and practice update. Symposium presented at the 29th Annual Conference of the Society for Industrial and Organizational Psychology, Honolulu, HI.*

This study was conducted in partnership with a client organization in the customer service field, with the objective of obtaining empirical support for the organization to deliver its pre-employment tests via mobile devices.

Illingworth, J., Morelli, N., Scott, S., & Boyd, S. (2014). Internet-based, unproctored assessments on mobile and non-mobile devices: Usage, measurement equivalence, and outcomes. *Journal of Business and Psychology, 30*, 25-34.

This study demonstrates that the use of mobile and non-mobile devices did not produce any practically significant score differences on assessment across

devices or applicant demographic subgroups.

King, D., Ryan, A. M., Kantrowitz, T. M., Grelle, G., & Dainis, A. (in press). Mobile Internet testing: An analysis of equivalence, individual differences, and reactions. *International Journal of Selection and Assessment, 4*, 382-394.

This paper examines the equivalence of mobile Internet testing (MIT) compared with testing on personal computers (PCs) and whether attitudes and other individual differences influence responses and reactions to MIT. Results demonstrated equivalence for a supervisory situational judgment test across testing modes, but not for a cognitive ability test. Respondents reported significantly more positive reactions when tested on a PC versus a mobile device.

Kinney, T. B., Lawrence, A., & Change, L. (2014). Understanding the mobile experience. Data across device and industry. In T. Kantrowitz & C. M. Reddock, *Shaping the future of mobile assessment: Research and practice update. Symposium presented at the 29th Annual Conference of the Society for Industrial and Organizational Psychology, Honolulu, HI.*

This study examined measurement equivalence, pass rate differences and applicant reactions by device type (smartphone, tablet, PC) for several mobile assessments; investigated the factors in mobile users' test environment; and conducted analyses across

three industries—manufacturing, customer service/retail and health care.

Landers, R. N., Reddock, C. M., Cavanaugh, K. J., & Proaps, A. J. (2014). Talent assessment using mobile devices. In T. Kantrowitz & C. M. Reddock, *Shaping the future of mobile assessment: Research and practice update. Symposium presented at the 29th Annual Conference of the Society for Industrial and Organizational Psychology, Honolulu, HI.*

The study consisted of 219 individuals who completed five commercially available and previously validated assessments from a large human resource consulting firm across two time points three weeks apart. The authors found relatively small differences in magnitude that were limited to cognitive-type measures.

Morelli, N. A., Mahan, R. P., & Illingworth A. J. (2014). Establishing the measurement equivalence of online selection assessments delivered on mobile versus non-mobile devices. *International Journal of Selection and Assessment, 22*, 124-138.

This study examines the measurement equivalence of selection assessments delivered on mobile and non-mobile devices. Measurement invariance tests conducted with multigroup confirmatory factor analysis suggest that mobile versions of a cognitive ability-type assessment, two bio data assessments, a multimedia work simulation and a text-based situational judgment

test appear to be equivalent to non-mobile versions. However, mobile device user latent means were half a standard deviation lower than their non-mobile counterparts for the situational judgment test.

Tippins, N. T. (2009). Internet alternatives to traditional proctored testing: Where are we now? *Industrial and Organizational Psychology: Perspectives on Science and Practice, 2, 2-10.*

This focal article explores Internet alternatives to traditional, proctored testing. It provides a comprehensive review of the definition and history of UIT, the factors that contribute to the rise of UIT and the mitigating factors that pose risks to organizations that use UIT.

COGNITIVE ABILITY TESTS

Gottredson, L. S. (1986). Societal consequences of the g factor in employment. *Journal of Vocational Behavior, 29, 379-410.*

This article discusses and rebuts seven common arguments that general intelligence (g) is of little or no practical importance in employment and illustrates the effect that differences in intelligence in a workforce may have on the structure and functioning of whole societies.

Hunter, J. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior, 29, 340-362.*

A review of the research literature indicates that general cognitive ability (GCA) predicts supervisor ratings and training success as well as objective, rigorously content-valid work sample performance. The author argues that the findings from this study support classic learning theory over behaviorist theories of learning and performance.

Hunter, J., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96, 72-98.*

An examination of research on predictors of job performance showed that for entry-level jobs cognitive ability was the most valid predictor. For selection on the basis of current job performance, the work sample test was slightly better. The authors suggested that using other predictors in conjunction with ability tests might improve validity and reduce adverse impact.

Ree, M. J., Earles, J. A., & Teachout, M. S. (1994). Predicting job performance: Not much more than g. *Journal of Applied Psychology, 79, 518-524.*

The authors investigated the roles of general cognitive ability (g) and specific abilities or knowledge (s) as predictors of work sample job performance criteria in seven jobs for U.S. Air Force enlistees. Analyses reveal that g was the best predictor and that s added a statistically significant but practically small amount to this prediction.

PERSONALITY ASSESSMENTS

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 91, 1-26.*

This study examines the validity of five personality measures for predicting performance in five occupational groups (professionals, police, managers, sales, and skilled or semi-skilled). Three different kinds of performance measures were used in this study: job proficiency, training proficiency and personnel data. One of the personality measures—conscientiousness—was consistently related to all three performance measures for all of the occupational groups. Extraversion was a valid predictor for two occupations involving social interaction—managers and sales. Openness to experience and extraversion were valid predictors of the training proficiency criterion. Overall, the results illustrate the benefits of using the five-factor model of personality in assessment research.

Costa, P. T., Jr., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13, 653-665.*

A considerable number of research studies have shown support for a five-factor model of personality traits. These include conscientiousness, extraversion, agreeableness, openness to experience and emotional stability. The claim that these factors represent the basic dimensions of personality is

based on four lines of reasoning: Longitudinal and cross-observer studies have demonstrated that all five factors are enduring dispositions that are manifested in patterns of behavior; traits related to each of the factors have been found in a variety of personality systems and in the natural language of trait description; the factors have been found in different age, sex, race and language groups, although they may be somewhat differently expressed in different cultures; and evidence of heritability has suggested that all have some biological basis.

Hough, L. M. (1992). The big five personality variables—construct confusion: Description versus prediction. *Human Performance*, 5, 135-155.

This article suggests that the five-factor model of personality traits (extraversion, conscientiousness, agreeableness, emotional stability and openness to experience) is not an adequate taxonomy of personality variables for predicting important job performance measures. The author argues that the five factors are too heterogeneous and incomplete. This paper first discusses the development of personality and job performance models. Comparisons between different taxonomies are made, and the author presents correlational evidence demonstrating the usefulness of nine personality measures (affiliation, potency, achievement, dependability, adjustment, agreeableness, intelligence, rugged individualism and locus of control).

Raymark, P. H., Schmit, M. J., & Guion, R. M. (1997). Identifying potentially useful personality constructs for employee selection. *Personnel Psychology*, 50, 723-736.

This article describes the Personality-Related Position Requirements Form (PPRF), a job analysis form used in making hypotheses about what personality predictors will be relevant for predicting performance in different jobs. The Big Five personality factors provided an organizing framework for the PPRF. Subsequent development resulted in identifying 12 specific sets of items for facets of each of the Big Five.

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, 44, 703-742.

This study examines past research to a) assess the overall validity of personality measures as predictors of job performance; b) investigate moderating effects of several study characteristics on personality scale validity; and c) investigate the predictability of job performance as a function of eight personality measures: neuroticism, extraversion, openness, agreeableness, conscientiousness, locus of control, Type A and miscellaneous. Consistent with predictions, studies using confirmatory research strategies produced corrected mean personality scale validity that was more than twice as high as that based on studies adopting exploratory strategies.

An even higher mean validity was obtained based on studies using job analysis explicitly in selection of personality measures.

BIO DATA TESTS

Hough, L. M. (1984). Development and evaluation of the “accomplishment record” methods of selecting and promoting professionals. *Journal of Applied Psychology*, 69, 135-146.

This study describes the development of the “accomplishment record” method, which resulted in self-reported descriptions of accomplishments in areas that were highly relevant for making selection or promotion decisions. The accomplishments were reliably rated using specially prepared rating scales and guidelines. To test the validity of the accomplishment record method, 329 attorneys prepared accomplishment records according to specified instructions and also completed a battery of more typical assessment methods. Scores on the accomplishment record inventory were unrelated to the traditional assessment measures (e.g., aptitude tests, grades and honors), but they correlated with job performance. The accomplishment record had also been found to correlate with self-perceptions of success, hard work and self-assurance and with the length of time spent practicing a profession.

Hough, L. M., Keyes, M. A., & Dunnette, M. D. (1983). An evaluation of three ‘alternative’ selection measures. *Personnel Psychology*, 36, 261-276.

The authors used a content-oriented validation strategy to develop three alternative selection inventories that were designed to reflect the job content of positions held by attorneys employed with a large federal agency. These inventories and three traditional assessment methods were completed by 329 agency attorneys as part of a concurrent validation study. Criterion-related validities of two traditional inventories (a background inventory and an interest and opinion inventory) and one alternative inventory (an accomplishment record inventory) were statistically and practically significant.

Shoenfeldt, L. F. (1999). From dustbowl empiricism to rational constructs in bio data. *Human Resource Management Review*, 9, 147-167.

This paper presents an attempt to use different types of biographical data scales in a concurrent validation to predict service orientation. Several types of scales showed positive validity results with the performance measures. The results were promising in terms of both the prediction and understanding of customer service orientation.

Stokes, G. S., Mumford, M. D., & Owens, W. A. (Eds.). (1994). *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction*. Palo Alto, CA: CPP Books.

This book is a comprehensive reference that provides information for understanding and using biographical data in selection in

the public and private sectors (including selection of blue-collar and federal government employees and work team members), career counseling and development, and job classification.

INTEGRITY TESTS

Camara, W. J., & Schneider, D. L. (1994). Integrity tests: Facts and unresolved issues. *American Psychologist*, 49, 112-119.

This article describes two independent reports, completed by the American Psychological Association (APA) and the U.S. Congress Office of Technology Assessment, that examined scientific and measurement issues concerning integrity testing. Data were presented on a variety of tests collected by surveying test publishers. The survey data provided a view of the industry's scope (e.g., test audience, user screening, score reporting) that was not available elsewhere. The article also addresses unresolved issues regarding integrity testing.

Goldberg, L. R., Grenier, R. M., Guion, L. B., Sechrest, L. B., & Wing, H. (1991). *Questions used in the prediction of trustworthiness in pre-employment selection decisions: An APA Task Force report*. Washington, DC: American Psychological Association.

This report presents the findings of an APA task force that was appointed to review available commercial tests used for the purpose of assessing the honesty and integrity of prospective employees. The report focuses exclusively on commercially published instruments

used for pre-employment selection decisions about the trustworthiness of job applicants.

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance [Monograph]. *Journal of Applied Psychology*, 78, 679-703.

The authors conducted a comprehensive examination of the validity of integrity tests based on 665 validity coefficients. Results indicated that integrity test validities were substantial for predicting job performance and counterproductive behaviors on the job, such as theft, disciplinary problems and absenteeism. Results of validity studies conducted on applicants using external performance measures (i.e., excluding self-reports) indicated that integrity tests predicted the broad criterion of organizationally disruptive behaviors better than they predicted employee theft alone. The authors concluded that integrity test validities were positive across situations and settings.

Sackett, P. R., & Wanek, J. E. (1996). New developments in the use of measures of honesty, integrity, conscientiousness, dependability, trustworthiness, and reliability for personnel selection. *Personnel Psychology*, 49, 787-829.

This article reviews journal articles, books, book chapters, law review articles, convention papers and dissertations regarding integrity testing for personnel selection.

Developments include an examination of professional and congressional inquiry into this area of testing, rapid growth of the validity database, new insight into similarities and differences between different tests, and links to the Big Five personality dimensions. Inquiries into relationships with other constructs are reviewed, as are applicant reactions to these tests.

STRUCTURED INTERVIEWS

Campion, M. A., Palmer, D. K., & Campion, J. E. (1997). A review of structure in the selection interview. *Personnel Psychology, 50*, 655-702.

This article reviews the research literature on the many ways selection interviews can be structured. Fifteen components of structure were identified that may enhance either the content or the evaluation process of the interview. The authors conclude that interviews could be easily enhanced by using the various components of structure. They recommend that improvement of this popular selection procedure be a high priority for future research and practice.

Campion, M. A., Pursell, E. D., & Brown, B. K. (1988). Structured interviewing: Raising the psychometric properties of the employment interview. *Personnel Psychology, 41*, 25-42.

This article proposes an employee-interviewing technique that includes the following steps: 1) develop questions based on a job analysis; 2) ask the same questions of each candidate; 3) use examples and illustrations;

4) have a panel record and rate answers; 5) administer the process to all candidates; and 6) emphasize job relatedness, fairness and documentation. When the interview was used to hire 149 entry-level production employees, it revealed high interrogator reliability and predictive validity as well as evidence for test fairness and utility.

Eder, R. W., & Ferris, G. R. (Eds.). (1989). *The employment interview: Theory, research, and practice*. Newbury Park, CA: Sage Publishing.

This book provides a comprehensive treatment of theory, research and practice relevant to the employment interview. The topics covered include the interviewer's decision-making process, applicant strategies and employment interview validity. There is also a concluding commentary summarizing the volume's implications for theory building, research methods and effective practice.

Judge, T. A., Higgins, C. A., & Cable, D. M. (2000). The employment interview: A review of recent research and recommendations for future research. *Human Resource Management Review, 10*, 383-406.

This comprehensive review examines research on the employment interview process. The authors began with a review of the traditional areas of interview research: reliability, validity, structured interviews, interviewer differences, equal employment opportunity (EEO) issues, impression management

and decision-making processes. Next, they discussed more recent developments in interview research, such as the use of the interview as a means of assessing person-organization fit and applicant reactions to the employment interview.

PHYSICAL FITNESS AND PHYSICAL ABILITIES TESTS

Blakley, B. R., Quinones, M. S., Crawford, M. S., & Jago, I. A. (1994). The validity of isometric strength tests. *Personnel Psychology, 47*, 247-274.

This article examines six studies in which isometric strength tests were used as part of selection procedures. The study focuses on the relationship between performance on four isometric strength tests and both supervisory ratings of physical performance and performance on work simulations. Results indicate that isometric strength tests were valid predictors of both types of performance measures across all of the jobs examined.

Campion, M. A. (1983). Personnel selection for physically demanding jobs: Review and recommendations. *Personnel Psychology, 36*, 527-550.

This review article argues that improvements in personnel selection systems for physically demanding jobs are needed due to EEO considerations, concern for worker physical well-being and the lack of alternative procedures.

Hogan, J. (1991). Physical abilities. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2, pp. 753-831). Palo Alto, CA: Consulting Psychologists Press.

This chapter identifies factors that affect the ability to perform in the workplace, including physiological responses, training, ability and injury, and workload. It reviews procedures and methods for designing personnel selection systems for physically demanding jobs and emphasizes the unique physical ability content of job analyses, performance measures and assessments. In addition, the chapter argues that the need to improve development and implementation of fair physical ability selection procedures and the need to understand the relationship between physical performance and other organizational effectiveness outcomes are important issues for future research.

SITUATIONAL JUDGMENT TESTS

Chan, D., & Schmitt, N. (1997). Video-based versus paper and pencil method of assessment in situational judgment tests: Subgroup differences in test performance and face validity perceptions. *Journal of Applied Psychology*, 82, 143-159.

This study examines the effects of race, reading comprehension, method of assessment, face validity perceptions and performance on a situational judgment test. The situational judgment test was administered via a videotape and via a paper-

and-pencil instrument. The results show that the racial differences in situational judgment test performance and face validity reactions to the test were substantially smaller with the video-based method of testing compared with the paper-and-pencil method.

Motowidlo, S. J., Dunnette, M. D., & Carter G. (1990). An alternative selection procedure: A low fidelity simulation. *Journal of Applied Psychology*, 75, 640-647.

From critical-incident analysis and judgments by subject-matter experts, a low-fidelity simulation was developed for selecting entry-level managers in the telecommunications industry. The simulation presented applicants with descriptions of work situations and five alternative responses for each situation. It asked them to select one response they would most likely make and one they would least likely make in addressing each situation. The results show that samples of hypothetical work behavior can predict performance, without the props, equipment or role players often required by high-fidelity simulations, such as work-sample tests or assessment centers.

Weichmann, D., Schmitt, N., & Harvey, V. S. (2001). Incremental validity of situational judgment tests. *Journal of Applied Psychology*, 86, 410-417.

Using three different samples, the authors assess the increases in the validity produced by situational judgment inventories (SJIs), relative to job knowledge,

cognitive ability, job experience and conscientiousness, in the prediction of job performance. The SJI was a valid predictor in three samples and produced substantial increases in validity in two samples. Relative to the other predictors, SJIs' correlation with performance, controlling for the other predictors, was superior in most comparisons. The authors concluded that the SJI should prove to be a valuable additional measure in the prediction of job performance, but they also suggested several additional areas of research.

WORK SAMPLE TESTS

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

This study reviews the validity evidence for a wide sample of motor and verbal work sample tests, which were designed to measure on-the-job behaviors. Motor tests were shown to have higher validities than verbal tests when job proficiency was the performance measure, but this pattern was reversed when success in training was the performance measure. These validities were somewhat lower than those for biographical information but higher than those for other predictors.

Gaugler, B. B., Rosenthal, D. B., Thornton, G. C., III, & Bentson, C. (1987). Meta-analyses of assessment center validity. *Journal of Applied Psychology*, 72, 493-511.

This paper examines the validity of assessment centers across

a large number of studies. Higher validities were found in studies in which ratings of potential were the performance measures, and lower validities were found in studies in which the assessment center was used to make promotion decisions. The authors also found that the validities were higher when a) the percentage of females who were being assessed was high, b) when several evaluation devices were used, (c) when assessors were psychologists rather than managers, d) when peer evaluation was used, and e) when the study was methodologically sound.

Howard, A. (1983). Work samples and simulations in competency evaluation. *Professional Psychology: Research and Practice*, 14, 780-796.

This article argues that the evaluation of professional competence in psychology—by licensing boards, the American Board of Professional Psychology or other boards—has not taken full advantage of expertise in psychological measurement within the profession. The advantages and disadvantages of work samples and simulations for such purposes are discussed and compared with those of paper-and-pencil tests. Examples from various professions and from psychological research in personnel selection are presented. A comparison of work samples with simulations shows greater advantages for the latter, but a combination of exercises in an assessment center model is recommended.

Thornton, G. C., III. (1992). *Assessment centers in human resources management*. Boston, MA: Addison-Wesley.

While assessment centers have been used for numerous human resource management functions, all assessment centers are not alike. This book shows how assessment for managerial selection/promotion, diagnosis of managerial training needs and team building among managerial groups are done differently. The author also shows how the assessed dimensions, observations and participant feedback all must be tailored to the specific application of the assessment center method.

Thornton, G. C., III, & Byham, W. C. (1982). *Assessment centers and managerial performance*. New York, NY: Academic Press.

This book focuses on examining the assessment center experience. It traces the historical development of multiple assessment procedures with emphasis on those advances relevant to assessment centers; critiques all of the published and unpublished research on assessment centers; integrates assessment center procedures into several theories of measurement and human judgment; and presents new models of job analysis, the nature of managerial work, work-sampling assessment methods and the process of human judgment based on the assessment center experience.

ADVERSE IMPACT

Roth, P. L., Bevier, C. A., Bobko, P., Switzer, F. S., & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology*, 54, 297-330.

The authors conducted a meta-analytic review on factors underlying the one standard deviation effect size difference observed in previous research for Black-White differences for college application tests (e.g., SAT) and overall analyses of tests of g for job applicants in corporate settings. They determined that the one standard deviation summary of group differences fails to capture many of the complexities in estimating ethnic group differences in employment settings.

Sackett, P. R., & Ellingson, J. E. (1997). The effects of forming multi-predictor composites on group differences and adverse impact. *Personnel Psychology*, 50, 707-721.

It is commonly assumed that including predictors that demonstrate smaller group differences with others that demonstrate larger group differences will help alleviate the adverse impact observed. The purpose of this paper was to answer the question, "If two or more predictors that have smaller and larger group differences are combined, what will be the magnitude of group differences and, consequently, of adverse impact?" A number of conclusions were drawn that clarified the extent to which combining

predictors with smaller and larger group differences affected subsequent adverse impact.

Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. (2001). High stakes testing in employment, credentialing, and higher education: Prospects in a post affirmative-action world. *American Psychologist*, 56, 302-318.

Cognitively loaded tests of knowledge, skill and ability often contribute to decisions regarding education, jobs, licensure or certification. Users of such tests often face difficult choices when trying to maximize both the performance and ethnic diversity of chosen individuals. The authors describe the nature of this quandary, review research on different strategies to address it and recommend using selection materials that assess the full range of relevant attributes using a format that minimizes verbal content as much as is consistent with the outcome one is trying to achieve. Regardless of the strategy adopted, however, they suggest that it is unreasonable to expect that one can maximize both the performance and ethnic diversity of selected individuals.

Sackett, P. R., & Roth, L. (1996). Multi-stage selection strategies. A Monte Carlo investigation of effects on performance and minority hiring. *Personnel Psychology*, 49, 549-562.

This study examines the effects of selection process variables on employee performance and on minority hiring. The authors look at how selection rules that did and did not include within-

group norming fared in terms of tradeoffs between performance and minority representation. The results showed that the preferred strategy depended on the relative value the organization placed on performance versus minority representation. Results also showed that the effects of different screen-then-select selection strategies varied as a result of the selection ratios at the screening and selection stages, thus precluding simple conclusions about the merits of each selection strategy.

APPLICANT REACTIONS

Bauer, T. N., Maertz, C. P., Dolen, M. R., & Campion, M. A. (1996). Longitudinal assessment of applicant reactions to employment testing and test outcome feedback. *Journal of Applied Psychology*, 83, 892-903.

This study uses a justice framework to examine reactions of candidates for an entry-level accounting job at three time periods: before testing, after testing but before feedback on whether they passed or failed the test, and after-test performance feedback. Several procedural justice measures (information known about the test, chance to perform, treatment at the test site, consistency of the test administration and job relatedness) predicted applicant evaluations regarding the organization, perceptions of employment testing and applicant test-taking self-efficacy. Test outcome favorability (passing or failing the employment test) predicted outcomes beyond initial

reactions more consistently than procedural justice perceptions. Procedural justice perceptions explained incremental variance in some analyses after the influence of outcome favorability was controlled.

Gilliland, S. W. (1995). Fairness from the applicant's perspective: Reactions to employee selection procedures. *International Journal of Selection and Assessment*, 3, 11-19.

Applicant reactions to selection procedures were examined in terms of the satisfaction and/or violation of 10 procedural justice rules. The author first collected 237 critical incidents describing fair and unfair treatment during selection processes from 31 individuals who had recently participated in job search and hiring. The critical incidents were categorized into 10 procedural justice rules, and the distribution of these incidents was examined for different hiring outcomes and different selection procedures. The most common procedural concerns reflected selection procedure job relatedness and the interpersonal treatment applicants had received.

Hausknecht, J. P., Day, D. V., & Thomas, S. C. (2004). Applicant reactions to selection procedures: An updated model and meta-analysis. *Personnel Psychology*, 57, 639-683.

An updated theoretical model of applicant reactions to selection procedures is proposed and tested in this article. Results indicate that applicants who held positive perceptions about selection were more likely to

view the organization favorably and report stronger intentions to accept job offers and recommend the employer to others. The average correlation between applicant perceptions and gender, age and ethnic background was near zero. Face validity and perceived predictive validity were strong predictors of many applicant perceptions, including procedural justice, distributive justice, attitudes toward tests and attitudes toward selection. Interviews and work samples were perceived more favorably than cognitive ability tests, which were perceived more favorably than personality inventories, honesty tests, bio data and graphology.

Ryan, A. M., Greguras, G. J., & Ployhart, R. E. (1996). Perceived job relatedness of physical ability testing for firefighters: Variations in reactions. *Human Performance, 9*, 219-240.

This research explores perceptions of physical abilities testing by examining the perceived job relatedness of physical abilities tests. Experience, self-efficacy and consistency of test administration were found to be related to perceptions of job relevance.

Smither, J. W., Reilly, R. R., Millsap, R. E., Pearlman, K., & Stoffey, R. (1993). Applicant reactions to selection procedures. *Personnel Psychology, 46*, 49-76.

The premise of this research is that applicant reactions to selection procedures may be of practical importance to organizations' attractiveness

to candidates and to selection procedure validity and utility. In part one of this two-part study, 110 newly hired entry-level managers and 44 recruiting-employment managers judged simulations, interviews and cognitive tests with relatively concrete item types (such as vocabulary, standard written English and mathematical word problems) as significantly more job-related than personality, bio data and cognitive tests with relatively abstract item types (such as quantitative comparisons and letter sets). A measure of new managers' cognitive abilities was positively correlated with the managers' perceptions of the job relatedness of selection procedures. In part two, the reactions of 460 applicants to a range of entry-level to professional civil service examinations were found to be positively related to procedural and distributive justice perceptions and willingness to recommend the employer to others.

AMERICANS WITH DISABILITIES ACT

Daley, L., Dolland, M., Kraft, J., Nester, M. A., & Schneider, R. (1988). *Employment testing of persons with disabling conditions*. Alexandria, VA: International Personnel Management Association.

This is a monograph based on a symposium presented at both the 1987 International Personnel Management Association Assessment Council (IPMAAC) Conference and the 1987 International Personnel Management Association (IMPA) International Conference. It

includes a review of research bearing on the accommodation of tests and testing procedures for examinees with disabilities and presents the experiences of two large eastern states and their programs of accommodation testing.

Eyde, L. D., Nester, M. A., Heaton, S. M., & Nelson, A. V. (1994). *Guide for administering written employment examinations to persons with disabilities*. Washington, DC: U.S. Office of Personnel Management.

This paper provides guidance for administering written examinations to persons with disabilities. It also includes information on legal requirements for reasonable accommodations and offers detailed guidelines for testing applicants who are deaf or hard of hearing or who have vision or motor impairments. The paper includes a checklist for physical accessibility of test sites and information on effective personal interactions with persons with disabilities.

Williams, S. K. (2003, January/February). *Tips for minimizing abuses under the Americans with Disabilities Act*. Alexandria, VA: Society for Human Resource Management.

This paper provides a discussion of the law, its interpretation and implications for use of various selection techniques in relation to the Americans with Disabilities Act of 1990. It covers physical exams, psychological exams, health questionnaires, physical agility tests, drug and alcohol tests, specific oral and

written questions (e.g., about the existence of a disability, about ability to perform job-related functions), and specific requests (e.g., how the applicant would perform job-related functions with or without reasonable accommodation).

LEGAL GUIDANCE

Azar, B. (1994). Could “policing” test use improve assessments? *APA Monitor*, 16.

This article discusses the American Psychological Association’s Standards for Educational and Psychological Testing. It discusses issues surrounding the enforcement of the standards as well as the possibility of developing more specific standards.

Society for Industrial and Organizational Psychology. (2003). *Principles for the validation and use of personnel selection procedures* (4th ed.). Bowling Green, OH: Author.

This document outlines principles adopted by the Society for Industrial and Organizational Psychology (SIOP) of the American Psychological Association for the validation and use of personnel selection and assessment procedures, including performance appraisal. The principles specify SIOP policy on conducting validation research using principles of good practice in the choice, development and evaluation of personnel selection procedures.

Uniform guidelines on employee selection procedures. (1978). *Federal Register*, 43, 38295-38315.

These guidelines incorporate a single set of principles designed to assist employers, labor organizations, employment agencies, and licensing and certification boards to comply with requirements of federal law prohibiting employment practices that discriminate on the grounds of race, color, religion, sex and national origin.

GENERAL REFERENCES

Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. Borman and Associates (Eds.), *Personnel selection in organizations* (pp. 71-98). San Francisco, CA: Jossey-Bass.

In this chapter, the authors argue that selection criteria should embrace a domain of organizational behavior broader than just technical task performance. Specifically, performance measures should be expanded to also include contextual activities, such as helping co-workers when needed and putting in extra effort to get the job done.

Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2014). Correlational effect size benchmarks. *Journal of Applied Psychology*, 100, 431-449.

The authors examine a large number (147,328) of correlations from published research in applied psychology reported

in two top-tier journals (*Journal of Applied Psychology* and *Personnel Psychology*) from 1980 to 2010 to produce empirical effect size benchmarks for 20 common research domains. Results indicate that the usual interpretation and classification of effect sizes in statistical methods books as small, medium and large bear almost no resemblance to findings in the field.

Cohen, J. (1988). *Statistical power for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

This book is a guide to power analysis in research planning that provides users of applied statistics with the tools they need for more effective analysis.

Finch, D. M., Edwards, B. D., & Wallace, J. C. (2009). Multistage selection strategies: Simulating the effects on adverse impact and expected performance for various predictor combinations. *Journal of Applied Psychology*, 94, 318-340.

The authors use Monte Carlo simulation to explore the trade-off between expected mean performance and minority hiring in multistage selection strategies and to identify those strategies most effective in balancing the trade-off. The results show that an increase in minority hiring was associated with a decrease in predicted performance for many scenarios, although the results also show that certain multistage strategies are much more effective than others for managing the performance and adverse impact trade-offs.

Heneman, H. G., III, & Judge, T. A. (in press). *Staffing organizations* (5th ed.). Middleton, WI: Mendota House, and New York, NY: McGraw-Hill.

The book is filled with up-to-date research, useful examples and best business practices in the staffing area. In addition, each chapter concludes with in-depth applications (cases and exercises) that enhance skills and provide practice in key staffing activities and related decision-making.

Kehoe, J. (2000). *Managing selection in changing organizations*. San Francisco, CA: Jossey-Bass.

This book provides managers and HR practitioners with practical guidance on making decisions about employee staffing. The chapters are authored by leading researchers and practitioners with extensive experience in the staffing area and offer proven strategies for the design and management of effective selection processes in organizations.

McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. A., & Ashworth, S. A. (1990). *Project A validation results: The relationships between predictor and criterion domains*. *Personnel Psychology*, 43, 335-353.

In this study, a predictor battery of cognitive ability, perceptual-psychomotor ability, temperament/personality, interest and job outcome preference measures was administered to enlisted soldiers in nine Army occupations. Scores from the cognitive and perceptual-psychomotor ability tests

provided the best prediction of job-specific and general technical task proficiency, whereas the temperament/personality measures best predicted giving extra effort, supporting peers and exhibiting personal discipline. Scores from the interest inventory correlated more highly with technical task proficiency than with demonstrating effort and peer support.

Murphy, K. (1996). *Individual differences and behavior in organizations*. San Francisco, CA: Jossey-Bass.

This book integrates existing research and stimulates new ways of thinking about how individual differences affect people's behaviors and experiences in organizations. It provides insights into what influences behavior on the job.

Potosky, D., & Bobko, P. (2004). *Selection testing via the Internet: Practical considerations and exploratory empirical findings*. *Personnel Psychology*, 57, 1003-1034.

This article presents equivalence information and practical lessons concerning selection testing via the Internet. The authors identify several issues associated with measurement and validity, the role of individual characteristics, respondents' reactions and behaviors, and other considerations regarding Internet test administration. They also report results from an exploratory study of the correlation between paper-and-pencil and Internet-administered cognitively oriented selection tests (including timed and untimed, proctored tests).

Ryan, A. M., & Tippins, N. T. (2004). *Attracting and selecting: What psychological research tells us*. *Human Resource Management*, 43, 305-318.

This article reviews research on which selection tools work, which recruitment strategies work, how selection-tool use relates to workforce diversity, and which staffing and recruiting processes lead to positive applicant perceptions. The goal is to assist the reader not only in understanding the present gaps between research and practice in recruitment and selection but also in developing skills for employing research in HR practice.

Rynes, S. L., Colbert, A. E., & Brown, K. G. (2002). *HR professionals' beliefs about effective human resources practices: Correspondence between research and practice*. *Human Resource Management*, 41, 149-174.

In this study, 5,000 human resource professionals were surveyed regarding the extent to which they agreed with various HR research findings. Responses suggested that large discrepancies existed between research findings and practitioners' beliefs in some areas, especially staffing. In particular, practitioners placed far less faith in intelligence and personality tests as predictors of employee performance than HR research would recommend.

Schmitt, N., & Borman, W. C. (1995). *Personnel selection in organizations*. San Francisco, CA: Jossey-Bass.

In this book, experts highlight the personnel selection issues that

will receive increasing attention in the years ahead. The book examines specific topics such as recruitment and retention, structured-versus-unstructured interviews, the ethics and effectiveness of computerized psychological testing, perceptions of selection fairness, productivity, turnover, and absenteeism.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.

This article summarizes the practical and theoretical implications of 85 years of research in personnel selection. It presents the validity of 19 selection procedures for predicting job and training performance and the validity of paired combinations of general mental ability (GMA) and the 18 other selection procedures.



Technology will continue to spark revolutionary changes in assessment methods and practices.





ENDNOTES

1. CEB. (2016). *Competing for critical talent* [Webinar].
2. CEB. (2012). Q4 global labor market survey.
3. CEB. (2010). Candidate rules of engagement survey.
4. CEB Corporate Leadership Council. *Creating enterprise leaders*.
5. Bauer, T. N., Maertz, C. P., Dolen, M. R., & Campion, M. A. (1996). Longitudinal assessment of applicant reactions to employment testing and test outcome feedback. *Journal of Applied Psychology, 83*, 892-903.
6. Kantrowitz, T. M. (2014). *Global assessment trends report 2014*. Arlington, VA: CEB.
7. Sullivan, J. (2014). The power has shifted to the candidate, so current recruiting practices will stop working. *ERE Recruiting Intelligence*. Retrieved from <http://www.ere-media.com/ere/the-power-has-shifted-to-the-candidate-so-current-recruiting-practices-will-stop-working/>
8. Beaty, J. C., Dawson, C. R., Fallaw, S. S., & Kantrowitz, T. M. (2009). Recovering the scientist-practitioner model: How IOs should respond to UIT. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 2*, 58-63.
9. Tippins, N. T. (2009). Internet alternatives to traditional proctored testing: Where are we now? *Industrial and Organizational Psychology: Perspectives on Science and Practice, 2*, 2-10
10. Ibid.
11. Golubovich, J., & Boyce, A. (2013). Hiring tests: Trends in mobile test usage. In N. Morelli, *Mobile devices in talent assessment: Where are we now?* Symposium presented at the 28th annual conference of the Society for Industrial and Organizational Psychology, Houston, TX.

12. Illingworth, J., Morelli, N., Scott, S., & Boyd, S. (2014). Internet-based, unproctored assessments on mobile and non-mobile devices: Usage, measurement equivalence, and outcomes. *Journal of Business and Psychology, 30*, 25-34.
13. Morelli N., Mahan, R. P., & Illingworth A. J. (2014). Establishing the measurement equivalence of online selection assessments delivered on mobile versus non-mobile devices. *International Journal of Selection and Assessment, 22*, 124-138.
14. Ibid.
15. Gael, S. (Ed.). (1988). *The job analysis handbook for business, industry, and government*. New York: Wiley; Brannick, M. T., & Levine, E. L. (2002). *Job analysis*. Thousand Oaks, CA: Sage.
16. LoVerde, M. A., & Schmidt, C. (2016, April). Overview of the CEB leadership study (CLS). In M. A. LoVerde, *Overview and selected findings from a multi-organizational, multi-level leadership study*. Symposium conducted at the 31st annual conference of the Society for Industrial and Organizational Psychology, Anaheim, CA.
17. Hunter, J. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior, 29*, 340-362.
18. Ree, M. J., Earles, J. A., & Teachout, M. S. (1994). Predicting job performance: Not much more than g. *Journal of Applied Psychology, 79*, 518-524; Gottfredson, L. S. (Ed.). (1982). The g factor in employment. *Journal of Vocational Behavior, 29*(3).
19. Raymark, M. J., Schmit, M. J., & Guion, R. M. (1997). Identifying potentially useful personality constructs for employee selection. *Personnel Psychology, 50*, 723-736.
20. Costa, P. T., Jr., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13*, 653-665.
21. Hough, L. M. (1992). The big five personality variables-construct confusion: Description versus prediction. *Human Performance, 5*, 135-155.
22. Tett, R. P., & Christiansen, N. D. (2007). Personality tests at the crossroads: A response to Morgeson, Campion, Dipboye, Hollenbeck, Murphy, and Schmitt (2007). *Personnel Psychology, 60*(4), 967-993.
23. Ibid.
24. Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology, 78*, 531-537.
25. Stokes, G. S., Mumford, M. D., & Owens, W. A. (Eds.). (1994). *Biodata handbook*. Palo Alto, CA: CPP Books.
26. Shoenfeldt, L. F. (1999). From dustbowl empiricism to rational constructs in biodata. *Human Resource Management Review, 9*, 147-167.
27. Hough, L. M. (1984). Development and evaluation of the "accomplishment record" methods of selecting and promoting professionals. *Journal of Applied Psychology, 69*, 135-146.
28. Hough, L. M., Keyes, M. A., & Dunnette, M. D. (1983). An evaluation of three 'alternative' selection measures. *Personnel Psychology, 36*, 261-276.
29. Hunter, J. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior, 29*, 340-362.
30. Motowidlo, S. J., Dunnette, M. D., & Carter G. (1990). An alternative selection procedure: A low fidelity simulation. *Journal of Applied Psychology, 75*, 640-647.
31. Kantrowitz, T. M. (2014). *Global assessment trends report 2014*. Alpharetta, GA: SHL.
32. Asher, J. J., & Sciarrino, J. A. (1974). Realistic work sample tests: A review. *Personnel Psychology, 27*, 519-533.
33. Hogan, J. (1991). Physical abilities. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2, pp. 753-831). Palo Alto, CA: Consulting Psychologists Press.
34. Hausknecht, J. P., Day, D. V., & Thomas, S. C. (2004). Applicant reactions to selection procedures: An updated model and meta-analysis. *Personnel Psychology, 57*, 639-683; Ryan, A. M., Greguras, G. J., & Ployhart, R. E. (1996). Perceived job relatedness of physical ability testing for firefighters:

- Variations in reactions. *Human Performance*, 9, 219-240; Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274; Roth, P. L., Bevier, C. A., Bobko, P., Switzer, F. S., & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology*, 54, 297-330; Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 91, 1-26.
35. Heneman, H. G., III, & Judge, T. A. (2006). *Staffing organizations* (5th ed.). Middleton, WI: Mendota House, and New York, NY: McGraw-Hill.
36. Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt, W. Borman, and Associates (Eds.), *Personnel selection in organizations* (pp. 71-98). San Francisco: Jossey-Bass; Hattrup, K., Rock, J., & Scalia, C. (1997). The effects of various conceptualizations of job performance on adverse impact, minority hiring, and predicted performance. *Journal of Applied Psychology*, 82, 656-664.
37. Sackett, P. R., & Roth, L. (1996). Multi-stage selection strategies. A Monte Carlo investigation of effects on performance and minority hiring. *Personnel Psychology*, 49, 549-562.
38. Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. (2001). High stakes testing in employment, credentialing, and higher education: Prospects in a post affirmative-action world. *American Psychologist*, 56, 302-318.
39. Ibid.
40. Heneman, H. G., III, & Judge, T. A. (2006). *Staffing organizations* (5th ed.). Middleton, WI: Mendota House, and New York, NY: McGraw-Hill.
41. Ibid.
42. Smither, J. W., Reilly, R. R., Millsap, R. E., Pearlman, K., & Stoffey, R. (1993). Applicant reactions to selection procedures. *Personnel Psychology*, 46, 49-76.
43. Gilliland, S. W. (1995). Fairness from the applicant's perspective: Reactions to employee selection procedures. *International Journal of Selection and Assessment*, 3, 11-19; CEB Recruiting Leadership Council. (2013). *Selecting high-quality hires for today's new work environment*; CEB Recruiting Leadership Council. (2013). *Candidates as customers: Optimizing the candidate experience*.
44. Kinney, T. B., Lawrence, A., & Change, L. (2014). Understanding the mobile experience: Data across device and industry. In T. Kantrowitz & C. M. Reddock, *Shaping the future of mobile assessment: Research and practice update*. Symposium presented at the 29th Annual Conference of the Society for Industrial and Organizational Psychology, Honolulu, HI.
45. Gutierrez, S. L., & Meyer, J. M. (2014). The mobile revolution: Measurement equivalence and mobile device administration. In T. Kantrowitz & C. M. Reddock, *Shaping the future of mobile assessment: Research and practice update*. Symposium presented at the 29th Annual Conference of the Society for Industrial and Organizational Psychology, Honolulu, HI.
46. Sireci, S. G., Scarpati, S. E., & Li, S. (2005). Test accommodations for students with disabilities: An analysis of the interaction hypothesis. *Review of Educational Research*, 75(4), 457-490.

The future of HR is here

(and it's SHRM-certified)



Today's workplace demands that HR professionals deliver innovative thinking and measurable business results. The SHRM-CP and SHRM-SCP, the new standard in HR certification, uses competencies to test not only what you know, but how you apply that knowledge in the workplace.

Join the more than 90,000 professionals who've earned their SHRM Certification, and demonstrate that you are the future of HR.



Get Certified Now:
shrmcertification.org/foundation



About the SHRM Foundation



At the SHRM Foundation, we are a catalyst for thought leadership. We help predict where the workforce is headed because we've been studying its evolution for nearly 50 years. Our mission is to offer unmatched human capital knowledge for the benefit of professional organization leaders with a total focus on studying and reporting the management practices that work. Supporting the Foundation is a chance to contribute to an ongoing study about the direction of human resources in society. The Foundation is governed by a volunteer board of directors, comprising distinguished HR academic and practice leaders. Contributions to the SHRM Foundation are tax-deductible. The SHRM Foundation is a 501(c)(3) nonprofit affiliate of the Society for Human Resource Management (SHRM). For more information, contact the SHRM Foundation at 703.535.6020. Online at shrmfoundation.org.

SHRM
Foundation
Shaping the future of HR