Evaluating Job Applicants:

The Role of Training and Experience in Hiring

A Report to the President and the Congress of the United States
by the U.S. Merit Systems Protection Board

JANUARY 2014

U.S. MERIT SYSTEMS PROTECTION BOARD
1615 M Street, NW
Washington, DC 20419-0001

The President
President of the Senate
Speaker of the House of Representatives

Dear Sirs:

In accordance with the requirements of 5 U.S.C. § 1204(a)(3), it is my honor to submit this U.S. Merit Systems Protection Board report, *Evaluating Job Applicants: The Role of Training and Experience in Hiring*. The purpose of this report is to help Federal agencies improve how they assess candidates for Federal jobs. Better qualified new hires mean higher productivity and better services to the public. Accurate assessments and unbiased selection procedures also support merit system principles such as selection and advancement based solely on relative ability, knowledge, and skills, fair and equitable treatment of applicants and employees, and efficient and effective use of the Federal workforce.

In times of fiscal austerity, it is essential that training and experience assessments play a central role in Federal hiring. Typically, training and experience assessments use past accomplishments as an indicator of an applicant's potential for job proficiency. While this approach is generally well regarded by applicants and managers, research shows that some of these assessments do little to predict successful job performance, and practical experience shows that it can be difficult to obtain accurate and detailed information from job applicants. This report examines how training and experience is assessed, identifies challenges in obtaining accurate and useful information from applicants, and provides strategies for improving such assessments.

As agencies become better at identifying and selecting the most productive workers, the Federal Government will operate more efficiently. I believe you will find this report useful as you consider issues affecting the Federal Government's ability to assess and select a high-quality workforce.

Respectfully,

Susan Tsui Grundmann

Enclosure

Evaluating Job Applicants:

The Role of Training and Experience in Hiring

Table of Contents

Executive Summary

Federal agencies are operating in a climate of fiscal austerity that has constrained budgets, increased the demands on current staff, and reduced opportunities for hiring. Although agencies must continue working towards mission accomplishment, filling vacant positions—and molding the workforce necessary for success—has become increasingly challenging. As such, it is essential that agencies and hiring managers make wise decisions about how to assess individuals' qualifications for any Federal jobs that they can fill.

This report discusses the advantages and disadvantages of using training and experience (T&E) assessments to gauge an individual's qualifications for a job and provides recommendations for the most effective use of such assessments.

Insights for this report were obtained from research and practice in the field of assessment, and from U.S. Merit Systems Protection Board (MSPB) survey data. This report adds value by synthesizing multiple perspectives on the advantages and challenges of T&E assessments and their implications for Federal hiring. We discuss these resources together in one document to demonstrate the role that T&E assessments can play in Federal hiring, highlight important themes and conclusions about the usefulness of T&E assessments, and assist Federal agencies and hiring managers in using T&E assessments appropriately.

Study Findings

Training and experience assessments are generally regarded as fair measures of job-related abilities and they are widely used by Federal agencies. Further, there are several T&E assessments that have practical value in Federal hiring. Training and experience assessments that are considered useful for predicting an applicant's future job performance include direct questions for factual information, occupational questionnaires, accomplishment records, and reference checks. Structured interviews and biodata, which can include this type of information, are also useful in gauging an applicant's likelihood of success in a job.

However, some of the T&E assessments that are commonly used in Federal hiring have poor ability to predict an applicant's future job performance. They include tallies of the years of job experience applicants have, the number of training classes they have attended, their grade point averages, and the contents of their resumes.

These measures are weak proxies for true assessment of an individual's capabilities, or present proficiencies, and should be avoided for hiring decisions.

Although there are several T&E assessments that can have value in Federal hiring, there are challenges with relying on an applicant's past to gauge his or her current proficiency or predict his or her future job success. First, trying to predict future job behavior from what we are told about the past is difficult because:

- What an applicant reports is not necessarily what happened;

- Applicants may not learn effectively from past events; and

- Training and experiences in the past may not be applicable to the present or future.

Second, there are challenges in how applicants assess their own proficiencies. Specifically:

- Applicants may have trouble recalling or reporting the most appropriate or applicable experiences;

- Applicants may not have an accurate perspective on the range of their proficiencies; and

- Applicants may not accurately evaluate their proficiencies.

Finally, applicants are not always honest in reporting or evaluating their past training or experiences.

Fortunately, our research shows there are strategies that can improve the usefulness of T&E assessments. These strategies focus on:

- Improving the accuracy of T&E assessments by grounding them in job analysis, and improving T&E questions, rating scales, scoring, and applicant self-assessment;

- Increasing the verification of applicant information by agency-driven approaches that rely on corroborating applicant claims, and through applicant-driven approaches such as warnings and requiring documentation or elaboration of claims;

- Using present-oriented assessments that examine the abilities and proficiencies that individuals currently possess; and

- Maintaining awareness of the latest developments in T&E assessment.

Conclusion

Training and experience assessments can bring value to a hiring or promotion process if implemented appropriately. There are advantages and disadvantages to using such assessments, and each hiring manager's threshold for the right balance will be different. Whenever possible, agencies should augment T&E measures by adopting one or more of the improvement strategies that are currently available. Further, agencies should remain abreast of developments in T&E assessment. Training and experience assessments, their improvement strategies, and future developments can play a critical role in shaping the effectiveness of an agency's overall assessment approach, workforce composition, and ultimately mission success.

Chapter 1 - Background

This study, which continues a series of studies focusing on Federal hiring and the assessment of job applicants[1], was conducted as part of the MSPB's responsibility to studies of Federal merit systems.[2] In its research, MSPB takes a Governmentwide perspective, looking across the practices of numerous employment contexts in Federal agencies with diverse missions and different cultures. The findings will be most useful to agency decision-makers when adapted to the context of their own workforce.

What is Assessment of Training and Experience?

This report examines how the previous learning and work-related experiences of job applicants are evaluated by Federal employers for hiring and promotion decisions. For the purposes of this report, *training* includes any structured experience undertaken for the primary purpose of gaining knowledge or skill in a specific area. Examples include training classes, college courses, webinars, individual study, and other similar activities. *Experience* includes not only on-the-job work tasks, but also any other activity that may convey job-relevant knowledge or skills or otherwise enhance one's ability to perform a particular job. In addition to official duties, experience may be gained from volunteer activities, recreational pursuits, or other non-work activities that develop abilities relevant to work.

Both training and experience can be evaluated, scored, and considered as factors in hiring and promotion decisions. The various tools for doing this are called ***training and experience (T&E) assessments***. The fundamental assumption behind all T&E assessments is that what we know in the present is shaped by what we have experienced in the past.[3] This assumption is one aspect of the principle of

---

[1] MSPB's series of assessment reports includes: MSPB (2003). The Federal Selection Interview: Unrealized Potential. MSPB: Washington, DC; MSPB (2005). Reference Checking in Federal Hiring: Making the Call. MSPB: Washington, DC; MSPB (2005). The Probationary Period: A Critical Assessment Opportunity. MSPB: Washington, DC; MSPB (2006). Reforming Federal Hiring: Beyond Faster and Cheaper. MSPB: Washington, DC; and MSPB (2009). Job simulations: Trying out for a Federal job. MSPB: Washington, DC.

[2] The Merit Systems Protection Board (MSPB) is an independent, quasi-judicial Federal agency that serves as the guardian of Federal merit systems. MSPB's responsibilities include conducting studies of Federal merit systems and the Federal workforce to ensure that Federal employees are managed in accordance with the merit system principles and in a manner free from prohibited personnel practices.

[3] Weisbuch, M., Slepian, M., Clarke, A., Ambady, N., and Veenstra-Vanderweele, J. (2011). Behavioral stability across time and situations: Nonverbal versus verbal consistency. *Journal of Nonverbal Behavior*, 34(1), 1-15.

behavioral consistency: the best predictor of future performance is past performance. Although this principle has proven reliable and useful across a variety of work tasks and settings, it has its limitations and exceptions.

We note that much of the information in this report is not new, *per se*. Certainly, readers can learn useful information about T&E assessments from a variety of sources. However, this report is a *synthesis* of multiple research and practitioner perspectives on the advantages and challenges of T&E assessments, and their implications for Federal hiring. Our goal is to demonstrate the role that T&E assessments can play in Federal hiring, highlight important themes and conclusions about T&E usefulness, and therein assist Federal agencies and hiring managers in using T&E assessments appropriately.

Advantages of training and experience assessments. Training and experience assessments prove advantageous when measuring the competencies of job applicants. For one, past experiences could appear to be more concrete than competencies, which must be inferred from a person's behaviors. Relying on past experiences avoids this inference because they appear to reflect actual competencies. In other words, people seem likely to be able to perform a job—or one similar to it—if they have done so in the past (experience). They also seem likely to be successful at a job if they have been told how to do it (training). For example, all applicants and hiring officials are familiar with resumes, which are basically lists of applicants' past experiences. Resumes can convey a wealth of information about an individual's qualifications for a job, and are particularly convenient for managers because much of the burden of preparation falls on applicants. There are advantages for applicants, too. Their job and training histories usually change slowly over time; once prepared, a resume need only be updated periodically. Applicants and hiring managers alike expect resumes to play a role in hiring. Most T&E assessments have at least a superficial resemblance to resumes, inheriting this high degree of acceptance as part of the hiring process. Commonly-used T&E assessments will be discussed in more detail in Chapter 2.

Disadvantages of training and experience assessments. Despite their allure, research has shown that there are challenges to using T&E measures. For example, many of these measures have relatively poor track records predicting future job performance. Some impose significant time burdens on applicants. Many are less effective at capturing skill levels accurately than measures that directly assess an individual's present proficiency or future potential. Training and experience measures that require applicants to evaluate their own experience present considerable opportunities for applicants to unintentionally (or intentionally) misrepresent their proficiency in an area. Simply, applicants may not be the best judges of their job-related qualifications. Even honest applicants with good intentions may not do this well. Evidence from both research and practice suggests that greater caution is appropriate

when using T&E measures to make important personnel management decisions. These challenges, possible strategies to address them, and general conclusions and recommendations on the use of T&E assessments are discussed in more detail later in this report.

Why Study Training and Experience Assessments?

MSPB is studying T&E assessments as part of its mission to uphold the Merit System Principles and support effective Federal workforce management. Additionally, T&E assessments are widely used and well regarded, and hiring decisions made using T&E assessments can have a lasting impact on the Federal workforce. Further, given the austere economic climate in which Federal agencies currently operate—and the resultant budget constraints and limited hiring opportunities—it is all the more critical that hiring managers make wise decisions about which assessments they use to bring in talent. As alluded to above, although T&E assessments can be useful in Federal hiring, they also have disadvantages that require attention and mitigation. Together, these factors make T&E assessments a prime candidate for study.

Merit system principles. Assessment and selection of Federal employees is central to MSPB's mission to promote the management of the civil service in accordance with the merit system principles.[4] The first principle requires that Federal employee "…selection and advancement should be determined solely on the basis of relative ability, knowledge and skills." When a selecting official considers prior training and experience as part of a hiring or promotion decision, the decision must be made that best identifies true differences in ability as accurately as possible. No assessment methods should be used which fail to meet professional testing standards for validity.[5]

The merit system principles further require that Federal employees be managed efficiently and effectively, as well as fairly and equitably. When assessments are not accurate in measuring abilities related to job performance, scores resulting from them are less reliable and the quality of decisions made from such scores is reduced. Such flawed decisions increase the likelihood of wasted resources, inefficient and ineffective practices, unfair personnel decisions, and an erosion of individual and organizational performance.

---

[4] Title 5 United States Code §2301 and §2302. The merit system principles and prohibited personnel practices referenced in this report can be accessed on the web at http://mspb.gov/meritsystemsprinciples.htm and http://mspb.gov/ppp/ppp.htm.

[5] Defined in the U.S. Office of Personnel Management's (OPM) *Assessment Decision Tool* as "The extent to which the assessment method has been shown to accurately measure a job-related competency and/or predict successful performance on the job." (Available at apps.opm.gov/ADT/content.aspx). The concept of validity will be discussed in more detail in Chapter 2, pg. 10.

Finally, it is also a merit system principle to educate and train employees when this will result in better organizational or individual performance. While this report does not focus directly on employee training, this principle is relevant because such training, sometimes provided by agencies, is often considered part of the evidence that an employee possesses certain proficiencies. This means that the relevance of training to the job, the quality of the training, and the degree to which training produces measurable learning collectively have an effect on the quality of a T&E assessment which seeks to measure it. When training is ineffective or evaluated ineffectively, error is introduced into the assessment of training and experience. Such a situation further reduces the value of T&E assessments as measures of job-relevant abilities.

Personnel assessment is an ongoing part of MSPB research because of the negative consequences a bad hire can have on an agency. The human resources (HR) community has long been aware of the damage a single individual can do through instability, incompetence, or dishonesty.[6] There are further costs in decreased morale and productivity while the rest of the organization struggles to cope with a bad hiring decision.[7] A previous MSPB study reported that the overall cost of hiring the wrong person for a job can be up to three times the employee's salary.[8] These costs can be magnified in the Federal workforce if an agency is slow to address a poorly-performing employee,[9] or supervisors are reluctant or unwilling to take corrective action,[10] permitting substandard performance to continue.[11]

Widely used, well-regarded. Training and experience assessments are used extensively in Federal hiring.[12] For example, "A 1999 MSPB study found that about 60 percent of delegated examining unit (DEU) hires are assessed through [ratings of] education and experience."[13] T&E is highly preferred as an assessment method when hiring Federal supervisors.[14] Further, "MSPB's 2000 Merit Principles Survey

---

[6] The costs associated with a bad hire are well known in the HR community; one organization's web site includes a "Bad Hire Calculator" that tallies them: www.adpselect-info.com/badHireCalculator.html.

[7] See, for example, Jay Goltz, "The Hidden Cost of Bad Hiring," The New York Times, March 1, 2011, available on the web at boss.blogs.nytimes.com/2011/03/01/the-hidden-costs-of-bad-hiring/.

[8] MSPB (2006). Reforming Federal Hiring: Beyond Faster and Cheaper. MSPB: Washington, DC.

[9] MSPB (2005). The Probationary Period: A Critical Assessment Opportunity. MSPB: Washington, DC.

[10] MSPB (2010). A Call to Action: Improving First-Level Supervision of Federal Employees. MSPB: Washington, DC.

[11] MSPB (2008). Attracting the Next Generation: A Look at Federal Entry-Level New Hires. MSPB: Washington, DC.

[12] See, for example, Sharpe, P. & Sorensen, K. (2011). Assessment of training and experience. IPAC 2011, Washington, DC. Also, MSPB (2003). The Federal Workforce for the 21st Century: Results of the Merit Principles Survey 2000. MSPB: Washington, DC; and Partnership for Public Service (2004). Asking the wrong questions: A look at how the Federal government assesses and selects its workforce. Partnership for Public Service: Washington, DC; and McDaniel, M., Schmidt, F. & Hunter, J. (1988). A meta-analysis of the validity of methods for rating training and experience in personnel selection. *Personnel Psychology*, 41, 283-314.

[13] MSPB (1999). The Role of Delegated Examining Units: Hiring New Employees in a Decentralized Civil Service. Washington, DC, p. 5.

[14] MSPB (2010). A Call to Action: Improving First-Level Supervision of Federal Employees. MSPB: Washington, DC.

(MPS) asked supervisors what information they use in selection decisions. Almost all—96 percent—reported using prior work experience to a great or moderate extent, and 82 percent reported using level of education to a great or moderate extent."[15] Thus, it appears that T&E measures are well-regarded by hiring managers.[16] A more recent survey of Federal HR practitioners found similar results. Measures of training and experience, such as resumes, occupational questionnaires, reference checks, Knowledge, Skills, and Abilities (KSA) narratives, and educational records, were among the most frequent assessments used in Federal hiring. In fact, the only non-T&E assessment among the six most frequent assessments was the structured interview.[17] Human resource practitioners report that training and experience measures continue to be widely used across public, private, and nonprofit sectors.[18] T&E assessments are also used across a wide range of situations, from entry-level assessment of basic qualifications to Executive Core Qualifications statements written by applicants to the Senior Executive Service.[19]

The preference for T&E assessments extends beyond Federal agencies. Job applicants may desire that all of their accomplishments be reviewed and evaluated. Some applicants may regard alternatives to this type of assessment, such as biodata and tests of cognitive ability, as biased or unfair. Training and experience, which acknowledge their work histories and achievements, may be more palatable. MSPB's study of fairness in Federal personnel practices notes that, "Education and training and length of experience received wide acceptance as they are likely viewed as fairly objective measures of an employee's ability to perform on the job."[20]

Yet, this regard for T&E assessments is at variance with research that identifies present-oriented assessments as more valid predictors of future job performance.[21] This report will discuss how such alternatives or focused improvements to T&E

---

[15] MPS 2000, Question 77A.

[16] For example, the 2000 Merit Principles Survey, which gathered information about the degree to which Federal supervisors value information from different types of assessment tools when making hiring decisions, found that T&E measures were well-regarded.

[17] MSPB Fair and Open Competition Survey (See http://mspb.gov/studies/surveys.htm).

[18] McGonigle, T. & Curnow, C. (2007). Measures of training and experience. In Applied Measurement: Industrial Psychology in Human Resources Management. Lawrence Erlbaum Associates: Mahwah, NJ.

[19] See www.opm.gov/policy-data-oversight/senior-executive-service/selection-process/#url=Process.

[20] MSPB (2013). Preserving the integrity of the Federal Merit Systems: Understanding and addressing perceptions of favoritism. MSPB: Washington, DC.

[21] Schmidt, F. & Hunter, J. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262–274.

measures can improve hiring across the Federal Government and how Federal agencies, Federal employees, and the public they serve can benefit.[22]

Lasting effects. Training and experience assessments can strongly shape who enters the Federal workforce. That is because such assessments, particularly those that can be implemented inexpensively, are most often used in the early stages of hiring as part of assessing minimal qualifications[23] or as an early screen in a series of assessments.[24] When properly documented with content validity evidence, education- and experience-based minimum qualifications have been upheld as valid and consistent with professional standards of assessment development.[25] However, it is particularly important that initial assessments be as accurate as possible. This is because decisions early in the hiring process that an applicant is "minimally qualified" may be difficult to reverse later in the process, even when a more rigorous assessment yields evidence that suggests an applicant has been misclassified.

Additionally, T&E assessments can affect an agency's workforce composition through unintentionally giving an advantage to internal applicants. This is a particular concern for jobs requiring the use of specific equipment or software that is not commonly used or available outside of an agency. Any internal employees who have had experience with such equipment or software would appear to have an immediate advantage over external candidates. Yet, both internal and external hires bring value to an agency. For example, internal hiring can build institutional knowledge and memory, and can provide employees with opportunities for growth and advancement, while external hiring can bring in new skills and fresh perspectives. The balance between internal and external hiring should be a conscious, strategic decision—rather than an unintended consequence of the criteria and methods used to assess job applicants.

Further, careless use (or misuse) of T&E assessments can have a lasting impact on an agency's reputation. It is a prohibited personnel practice to structure a job competition to advantage or disadvantage a particular individual. Training and experience assessments that focus too narrowly on a very specific set of training,

---

[22] T&E assessments have also been used extensively in state and local government hiring. A review of best practices will be useful to these efforts as well. See Johnson, J., Guffey, W. & Perry, R. (1980). When is a T and E rating valid? IPMAAC: Boston, MA;IPMA-HR (2006). Recruitment and Selection Benchmarking. Alexandria, VA.

[23] See OPM's Operating Manual for Qualification Standards for General Schedule (GS) Positions retrievable from www.opm.gov/qualifications. See also OPM's Job Qualification System for Trades and Labor Occupations retrievable from www.opm.gov/qualifications/x-118c.

[24] For example, see Agarwal, S., Busciglio, H., Leaman, J., Simmons-Collins, T., Curtin, P. & Thompson, S. (2010). Understanding the practical outcomes of self-report, task-based assessment questionnaires for assessing applicants for U.S. Customs and Border Protection vacancies. SIOP, 2010.

[25] Buster, M., Roth, P. & Bobko, P. (2009). A process for content validation of education and experience-based minimum qualifications: An approach resulting in Federal court approval. *Personnel Psychology*, 58(3), 583-630.

education, and experience factors can be seen as inappropriately narrowing the field of candidates, even if this was not the intent. As a result, some qualified individuals may self-select out of the application process, believing it to be a waste of their time.

Given the relevance of T&E assessments to the merit system principles, their widespread use, and their lasting effects on the Federal workforce, it is important that Federal agencies understand how such assessments can be used most effectively. Further, Federal budget constraints and limited hiring opportunities reinforce the importance of agencies and hiring managers making wise decisions about assessments and the hiring decisions that follow.

Study Methodology

Several sources of data and information were compiled for this report. These sources are of three general types: current best practices in measurement; Governmentwide surveys; and study-specific surveys conducted by MSPB.

Best practices. Information about the relative validity (or usefulness) of T&E measures was obtained from a review of the applied psychology literature as reported in professional journals and paper presentations at meetings of professionals who work in applied psychology. We have examined and referenced reports on T&E assessments in both public and private sectors to identify techniques which improve their effectiveness. Information about how T&E assessments are used by human resources (HR) practitioners was obtained from HR professional associations and through interviews with Federal HR and personnel specialists.

Governmentwide surveys. Data about the use of assessments in Federal agencies was gathered from surveys conducted by MSPB. These surveys are of two types. First, MSPB periodically conducts Governmentwide Merit Principles Surveys (MPSs) of a portion of the Federal workforce. These surveys are conducted every three to five years and collect responses from 40,000 to 50,000 participants. These surveys provide trend information about agency management practices and employee perceptions of their work environment. Several MPSs have focused on the use and value of assessments for selection and promotion. Details about sampling, administration, and results of these surveys are reported in the project reports for each survey.

Study-specific surveys. Second, MSPB also conducts study-specific surveys that focus on topics such as fairness, management practices, and human resources (HR). Several of these surveys provided data referenced in this report. Details of the design and administration of these surveys are reported in the project reports referenced in this report.

In This Report

As part of MSPB's study series on assessment, this report is written to help improve assessment practices in the Federal Government.[26] Assessment can be improved through strategies shown to increase their quality of T&E assessments or by using alternative assessments. MSPB offers research-based steps for improvement.

Four chapters follow this introduction (Chapter 1):

- Chapter 2 reviews the common methods that are often used to assess training and experience. These methods range from simplistic counting of years of job experience to in-depth analyses of employee accomplishments.

- Chapter 3 describes several factors that can reduce the ability of any T&E measure to predict job performance, including a focus on the past, unreliable applicant self-assessment, and applicant dishonesty.

- Chapter 4 suggests strategies to either improve the accuracy of existing T&E assessments, or replace T&E assessments with alternative methods.

- Chapter 5 outlines this report's conclusions and recommendations for using T&E assessment in the Federal workplace.

---

[26] MSPB's series of assessment reports includes: MSPB (2003). The Federal Selection Interview: Unrealized Potential. MSPB: Washington, DC; MSPB (2005). Reference Checking in Federal Hiring: Making the Call. MSPB: Washington, DC; MSPB (2005). The Probationary Period: A Critical Assessment Opportunity. MSPB: Washington, DC; MSPB (2006). Reforming Federal Hiring: Beyond Faster and Cheaper. MSPB: Washington, DC; and MSPB (2009). Job simulations: Trying out for a Federal job. MSPB: Washington, DC.

Chapter 2 - How is Training and Experience Assessed?

This chapter introduces common methods for assessing training and experience and summarizes how they are administered and how well they predict future job performance. The chapter examines T&E assessments and assessments using T&E information which predict job performance reasonably well and then a small number of problematic T&E assessments.

Determining Assessment Quality

MSPB has advised agencies to "[e]mploy rigorous assessment strategies that emphasize selection quality, not just cost and speed. In particular, use assessment instruments that have a relatively good ability to predict future performance."[27] The quality standards for assessments used for selection and promotion must be high for at least two reasons. First, decisions made using them have a great impact on the lives and careers of individuals and on the effectiveness of employing organizations. Second, these decisions must sometimes withstand legal challenge. Both considerations require that assessments conform to the best practices and standards of professional assessment development.

Sources of standards. Fortunately, assessment of human capabilities is a well-developed discipline. Through research and experience, the field has established a set of professional standards for producing high quality assessments. MSPB guidance is available from several assessment-related reports. MSPB has also produced a report, *Reforming Federal Hiring: Beyond Faster and Cheaper*, that summarizes its research and recommendations on assessment.

The Office of Personnel Management's *Assessment Decision Tool*[28] summarizes much of this information with an applied focus. It refers to the other sources listed below. If your reading time is limited, this is probably the best single source.

The Department of Labor's *Testing and Assessment: An Employer's Guide to Good Practices* (www.onetcenter.org/guides.html) is also written for practitioners, but requires some background in personnel selection.

---

[27] MSPB (2006). Reforming Federal Hiring: Beyond Faster and Cheaper. MSPB: Washington, DC.

[28] Available at apps.opm.gov/ADT/content.aspx.

The *Uniform Guidelines on Employee Selection Procedures* (www.uniformguidelines. com), contained in 29 CFR Part 1607, are somewhat more comprehensive.

Professional guidance intended for all sectors can be found in two documents. The *Standards for Educational and Psychological Testing* (www.apa.org/science/standards. html) was produced collaboratively by the American Psychological Association (APA), the National Council for Measurement in Education (NCME), and the American Educational Research Association (AERA). Finally, the *Principles for the Validation and Use of Personnel Selection Procedures* (www.siop.org/_Principles/ principlesdefault.aspx) is maintained by the Society for Industrial-Organizational Psychology (SIOP), the primary professional association for those who focus on personnel selection.

Validity. In accordance with the Merit System Principles hiring managers are responsible for selecting individuals for jobs based on their relative merits, knowledge, skills, and abilities. Yet, it can be challenging to gauge an individual's merits for a particular job, or evaluate how likely it is that an individual will be a good performer in that job. Fortunately, as discussed above, there is a professional discipline (and standards) devoted to developing quality assessments of individuals' knowledge, skills, and abilities, and using such assessments to predict job performance. Also fortunate is that there is a standard for determining how good, or useful, an assessment is for a particular job. This standard is called "validity." OPM's *Assessment Decision Tool* defines validity as, "The extent to which the assessment method has been shown to accurately measure a job-related competency and/or predict successful performance on the job."[29]

Validity can be reflected in a number that is called the "validity coefficient." In essence, the validity coefficient is a rule of thumb for gauging assessment usefulness for predicting how someone will perform on a particular job. Just as an individual's credit score is a number used to reflect how likely a person is to be able to pay back a loan, an assessment's validity coefficient is a number that can be used to reflect the likelihood that an applicant's scores on a particular assessment can be used to predict their future job performance. The process of determining a validity coefficient is beyond the scope of this study.[30] For current purposes and subsequent chapters, it is sufficient to keep in mind the guidelines in Table 1.

---

[29] *Assessment Decision Tool*, OPM.

[30] See Appendix A for a further discussion of validity and validity coefficients.

Table 1: Range of Values for Validity Coefficients[31]

Validity Coefficient Value	Usefulness for Predicting Job Performance
.35 to 1.0	Very Useful
.21 to .35	Likely to be Useful
.11 to .21	Probably Useful
.11 or below	Unlikely to be Useful

It is also important to keep in mind that the validity coefficient can never predict perfectly who is (or is not) likely to succeed in a job.[32] For example, a person's ability as measured at the time they are hired is only one factor that influences their future job performance. A person's future job performance will also be affected by their motivation, engagement, quality of supervision, organizational support of performance, and other contextual factors. Measuring an individual's performance is also contaminated by factors such as favoritism, performance rating inflation, and idiosyncratic performance standards.

Agencies and practitioners should consider the validity coefficients of T&E assessments when deciding which are most appropriate for each hiring decision. Organizational effectiveness research demonstrates that good personnel selection procedures, made possible through the use of high quality assessments, contribute measurably to organizational performance.[33] We will place each type of assessment into one of the categories of usefulness in Table 1.[34] Such evidence can assist agencies and hiring managers in understanding the usefulness of a particular assessment for predicting applicant's likelihood of success on a job. In light of the above discussion of assessment validity, the following sections provide an overview of common T&E assessments. The goal is to provide a general understanding of the overall form and usefulness of each assessment for identifying applicants who are likely to be successful in particular jobs. These assessments

---

[31] Sproule, C. (2009). Rationale and research evidence supporting the use of content validation in personnel assessment. IPAC. Available at www.ipacweb.org.

[32] In other words, we cannot expect validity coefficients to approach 1.0 for many reasons which have nothing to do with assessment quality. *See* Schmidt, F., Caplan J., Bemis, S., Decuir, R., Dunn, L. & Antone, L. (1979). The behavioral consistency method of unassembled examining (TM-79-21). Washington, DC: U.S. Office of Personnel Management, Personnel Research and Development Center.

[33] Van Iddekinge, C., Ferris, G., Perryman, A., Blass, F. & Heetderks, T. (2009). Effects of selection and training on unit-level performance over time: A latent growth modeling approach. *Journal of Applied Psychology*, 94(4), 829-843.

[34] Appendix B contains the numeric validity coefficients used to place each specific type of T&E assessment into the appropriate category in Table 1.

may be used in a variety of contexts—to evaluate minimum qualifications, measure entry level skills, distinguish between candidates later in the assessment process, or determine eligibility for promotion.

Useful T&E Assessments

T&E assessments that have proved useful in predicting future job performance include direct questions, occupational questionnaires, accomplishment records, KSA narratives, and reference checks.

Direct questions. Applications for Federal jobs can ask for factual information, such as college major, time in grade, credit hours completed, degrees or certifications held, past or present employment with the Federal government or with a particular agency, veteran's status, or whether the applicant meets some standard. Response options for such questions are typically "Yes/No" or multiple-choice. This approach is often encountered in assessing minimum qualifications.[35] For example: "This GS-14 position requires that you have at least one year of service at the GS-13 level or equivalent. Do you meet this requirement?" Or, "This position requires that you have earned at least 12 credits in computer science or a related field at an accredited college or university. Do you meet this requirement?"

Given the wide range of direct questions that could be asked, we cannot report one validity coefficient for this method. Yet, the usefulness will depend on whether the information requested has been shown by job analysis to be related to success on the job, whether the applicant supplies the requested information, and whether the information is confirmed to be accurate.

As an aside, it is a good idea to list such questions in a vacancy announcement so applicants can read through them and screen themselves out if they do not meet the criteria described. This kind of self-selection is good for everyone—it reduces the assessment burden on the agency to process applications from unqualified applicants and reduces wasted effort on the part of applicants as well. However, this only works if the question's meaning is clear and seeks information that is verifiable. For example, it would not be a good idea to present a densely-written paragraph of job qualifications and ask applicants: "Do you meet these qualifications?" Not only is this overwhelming, but it requires applicants to make judgments they may not know how to make.

Occupational questionnaires. Occupational questionnaires bridge the "Possibly Useful" and "Likely Useful" categories for predicting future job performance (see Table 1). Validity coefficients for occupational questionnaires have ranged from .15 to

---

[35] *OPM Delegated Examining Operations Handbook* available at www.opm.gov/deu/.

.28, making the quality of the particular questionnaire a crucial factor in determining its usefulness.[36] A typical occupational questionnaire presents a list of specific job tasks or behaviors and asks applicants to rate their level of experience performing each. The tasks (items) are often developed to measure specific competencies or KSAs. They are job related and generally easy for experienced workers to rate. Different questionnaires vary in what information they ask applicants to provide about each task. Common strategies are to ask applicants whether or not they have ever performed the task, how often they have performed it, how much time they have spent performing it, how effectively they have performed it, whether or not they received training on how to perform it, how closely they were supervised, and whether they have ever taught it to others.[37]

Occupational questionnaires have been used extensively to quantify an applicant's degree of experience with specific tasks performed or abilities (competencies) required on the job and likely performance on a similar job in the future.[38] In 2006, one study found that 70 percent of public sector organizations (local, state and Federal) reported using questionnaire-based T&E ratings as one of their three most frequently-used assessments.[39] More recent discussions note that such questionnaires are used extensively by Federal agencies, particularly those employing automated staffing systems.[40] Their use has increased since hiring reform guidelines issued by the U.S. Office of Personnel Management[41] have removed KSA narratives and similar written assessments from the initial screening of applicants.

Occupational questionnaires are best used to screen out poor candidates rather than to distinguish between the best candidates. In particular, they are most helpful in screening a large applicant pool representing a broad range of differences in ability

---

[36] Schmidt, F. & Hunter, J. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262–274; and Schneider, R. (1984). The rating of training and experience: A review of the literature and recommendations on use of alternative E&E procedures. *Personal Assessment Monographs*, 3(1), IPMAAC: Alexandria, VA.; and McDaniel, M., Schmidt, F. & Hunter, J. (1988). A meta-analysis of the validity of methods for rating training and experience in personnel selection. *Personnel Psychology*, 41, 283-314.

[37] McGonigle, T. & Curnow, C. (2007). Measures of training and experience. In Applied Measurement: Industrial Psychology in Human Resources Management. Lawrence Erlbaum Associates: Mahwah, NJ.

[38] Quinones, M., Ford, K. & Teachout, M. (1995). The relationship between work experience and job performance: A conceptual and meta-analytic review. *Personnel Psychology*, 48, 887-910.

[39] IPMA-HR (2006). Recruitment and Selection Benchmarking. Alexandria, VA.

[40] Barton, M. (2010). Automated T&E questionnaires: Practical outcomes and development considerations. SIOP.

[41] See www.opm.gov/policy-data-oversight/human-capital-management/hiring-reform/.

rather than in assessing a more experienced pool of well qualified candidates.[42] They are therefore useful in the first stage of multiple hurdle assessment plan.[43]

Applicant ratings on occupational questionnaires are scored using question and/or competency weighting strategies developed during job analysis. Some questions or competencies may be weighted more heavily depending on the criticality of the tasks/competencies to the job, or abilities that are difficult to acquire or rarely found in the applicant pool. But the scores also depend on the judgment of the applicants themselves as they decide how to translate their experiences into the language and scales of the questionnaire.

Over time, research and practice have made improvements to occupational questionnaires. There is some evidence that future performance can be better predicted when questions address how frequently tasks are performed rather than the total amount of time spent performing them. Predictive power also increases when assessments ask about more specific tasks or competencies that are closely job-related than about general types of tasks or competencies.[44] They are more effective for predicting performance in professional/administrative/scientific positions than in clerical/technical positions.[45]

The U.S. Office of Personnel Management (OPM) has a long history of using occupational questionnaires, ranging from the initial use of scanned paper forms to the current web-based questionnaires.[46] OPM has worked with other agencies to refine and disseminate best practices for using these assessments.

Accomplishment records. Accomplishment records are in the "Very Useful" category for predicting future job performance (see Table 1). Accomplishment records achieve

---

[42] Agarwal, S., Busciglio, H., Leaman, J., Simmons-Collins, T., Curtin, P. & Thompson, S. (2010). Understanding the practical outcomes of self-report, task-based assessment questionnaires for assessing applicants for U.S. Customs and Border Protection vacancies. SIOP, 2010; Barton, M. (2010). Automated T&E questionnaires: Practical outcomes and development considerations. SIOP; and McDaniel, M., Schmidt, F. & Hunter, J. (1988). A meta-analysis of the validity of methods for rating training and experience in personnel selection. *Personnel Psychology*, 41, 283-314.)

[43]Sproule, C. (2009). Rationale and research evidence supporting the use of content validation in personnel assessment. IPAC. (Available at www.ipacweb.org.)

[44] Quinones, M., Ford, K. & Teachout, M. (1995). The relationship between work experience and job performance: A conceptual and meta-analytic review. *Personnel Psychology*, 48, 887-910.

[45] Agarwal, S., Busciglio, H., Leaman, J., Simmons-Collins, T., Curtin, P. & Thompson, S. (2010). Understanding the practical outcomes of self-report, task-based assessment questionnaires for assessing applicants for U.S. Customs and Border Protection vacancies. SIOP; and Barton, M. (2010). Automated T&E questionnaires: Practical outcomes and development considerations. SIOP.

[46] OPM (2009). Developing and Administering Automated Task- and Competency-Based Questionnaires. Also Barton, M., Bisges, J. & Holloway-Lundy, A. (2010). Practical outcomes associated with use of automated questionnaires at the U.S. Office of Personnel Management. SIOP 2010.

validities of .45, higher than for any other T&E measure.[47] "The accomplishment record is a means of gathering self-reported and verifiable descriptions of experience on relevant behavioral job dimensions."[48] It was developed to collect information from professionals who sometimes do not respond favorably to more test-like assessments.[49] Applicants provide detailed, written descriptions of a small number of accomplishments that they choose to best illustrate their proficiency level on dimensions identified as important by a job analysis. These accomplishments need not be from the work setting, but may be from other areas of the applicant's life. Applicants can be asked to provide supplemental information about the context of the accomplishment, the steps taken to accomplish it, the resources used, their share of responsibility for the work, and the product or outcome.

One key to this method's high validity is that the information applicants provide is verifiable. They are asked to provide names and contact information for individuals who can verify that their description of the accomplishment is accurate. Another important factor is that scoring is based on information about successful job performance gathered during job analysis.

However, accomplishment records are expensive to develop. The accomplishment record is sometimes used later in a multiple hurdle selection process as a verification method for an earlier occupational questionnaire because the scoring process for accomplishment records can be more labor intensive than these other methods.[50]

Past work samples, such as reports applicants present as their work, are a variation of the accomplishment record; there is no need to have the applicant write a description, since the product of the original work is available for evaluation. But this product is of uncertain parentage without further information about the applicant's role in producing it. And past work samples from different applicants may not be easy to compare. As with accomplishment records, contact information can be collected from the applicant and be used to independently verify the applicant's level of involvement in producing the work sample.

---

[47] Schmidt, F. & Hunter, J. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), p. 268.

[48] McGonigle, T. & Curnow, C. (2007). Measures of training and experience. In Applied Measurement: Industrial Psychology in Human Resources Management. Lawrence Erlbaum Associates: Mahwah, NJ.

[49] Hough, L. M. (1984). Development and evaluation of the "accomplishment record" method of selecting and promoting professionals. *Journal of Applied Psychology*, 69, 135-146.

[50] Current OPM hiring reform guidance also precludes the use of the accomplishment record and other assessments requiring written narratives as part of the initial screen in Federal hiring.

KSA narratives. The KSA narratives are another variation of the accomplishment record.[51] KSA narratives and accomplishment records are similar assessments, differing largely in the degree of structure in development, presentation, and scoring.[52] A KSA question presents a competency, job, or job task and asks applicants to write a narrative evidencing that they have this competency or can do this job or task. When a job is broken down into these pieces, it becomes easier for applicants to highlight their best example for each KSA instead of selecting a single accomplishment which may not demonstrate their full strength in every area critical to the job. KSA questions are often no more than a sentence or two, and applicants are encouraged to show brevity in their replies, which does not permit the in-depth analysis that a longer accomplishment record narrative typically permits. Applicants are typically not asked for a reference for each KSA, reducing the possibility that an official can later confirm the accuracy of the applicant's assertions.

Although the use of KSA narratives permits an agency to identify multiple applicant capabilities, KSA narratives require a significant time commitment from applicants to write and an equal or greater commitment from agencies to score. MSPB has long been concerned about this applicant burden: KSAs "can be time-intensive and burdensome to complete. In addition, they can be alien and discouraging to applicants because private sector organizations do not generally require these types of submissions."[53] This burden extends to selecting officials as well: "In a 2004 OPM survey, almost 50 percent of agency respondents said that a key barrier to timely hiring is the amount of time spent by selecting officials in reviewing applicant credentials and conducting interviews."[54]

In 2010, a significant portion of this burden was lifted. President Obama signed an Executive Order to remove KSA narratives from the initial stages of screening for Federal jobs. OPM Director John Berry praised this change for introducing "commonsense hiring" which would allow selections to proceed "efficiently and quickly."[55] However, in some agencies the burden has been eased only slightly, as KSA narratives were simply moved later in the hiring process. Other agencies dropped KSA narratives in favor of less burdensome occupational questionnaires for initial employment screening.

---

[51] No validity coefficient is available for KSA narratives.

[52] Bradley, K. (2011). Testing and Assessment Methods for Public Sector Hiring: Supplementing (or Supplanting?) T&E Evaluations. IPAC.

[53] MSPB (2006). Reforming Federal Hiring: Beyond Faster and Cheaper. MSPB: Washington, DC.

[54] MSPB (2006). Reforming Federal Hiring: Beyond Faster and Cheaper. MSPB: Washington, DC, which cites United States House of Representatives, Statement of the Honorable Dan G. Blair, Deputy Director, Office of Personnel Management, Before the Subcommittee on Civil Service and Agency Organization, Committee on Government Reform, Chicago, IL, June 7, 2004.

[55] "Remarks of OPM Director John Berry, Hiring Reform First Anniversary Event", May 18, 2011. Available at www.opm.gov/news/speeches-remarks/hiring-reform-first-anniversary-event/.

There is some indication that the impact of reducing applicant burden may not be entirely positive. Many HR professionals report an increase in application volume that is not necessarily accompanied by an increase in qualified applicants.[56] More individuals apply for Federal jobs who may be less serious about obtaining those jobs. Their application materials must still be processed and evaluated, increasing agency burden significantly.

Reference checks. Reference checks have an average validity coefficient of .26,[57] and are therefore in the "Likely Useful" category for predicting future job performance (see Table 1). A reference check is a structured discussion between a potential employer and those who are acquainted with an applicant's previous work and job-relevant behavior. OPM defines it similarly: "Reference checking is an objective evaluation of an applicant's past job performance based on information collected from key individuals (e.g., supervisors, peers, subordinates) who have known and worked with the applicant." An MSPB survey of Federal HR professionals found that reference checks play a role in nearly 80 percent of Federal hiring processes.[58] MSPB has discussed reference checking extensively in the report *Reference Checking in Federal Hiring: Making the Call.*[59]

Some T&E assessments suffer from a lack of verification. Reference checks are at the opposite end of the spectrum. They are *primarily* verification of previously-obtained information (often from a resume) using the applicant's professional network. Such verification allows reference checks to add to the usefulness of other assessment methods, such as resumes, which will be discussed below. One of the best features of reference checks is that they are typically performed at the end of the assessment process when management has narrowed the applicant pool to a very few job candidates. Thus, the investment of time is focused on only those individuals to whom the agency expects to make a job offer. The ability to confirm information given by candidates earlier in the process is an ample return on the investment of time in checking a few references.

Assessments that Include T&E Information

The T&E assessments discussed in the previous section all focus on gauging an individual's current proficiency from their experience gained and abilities developed at some time in the past. There are assessments which gather the same type of information about the past, but combine it with other information to produce a

---

[56] MSPB Fair and Open Competition Survey (See http://mspb.gov/studies/surveys.htm).

[57] Schmidt, F. & Hunter, J. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262–274.

[58] MSPB Fair and Open Competition Survey (See http://mspb.gov/studies/surveys.htm).

[59] MSPB (2005). Reference Checking in Federal Hiring: Making the Call. MSPB: Washington, DC.

score. This section briefly describes two such measures: structured interviews and biodata questionnaires.

Structured interviews. This assessment has a predictive validity of .51,[60] and is therefore in the "Very Useful" category for predicting future job performance (see Table 1).   Key characteristics of structured interviews include: each applicant is asked the same set of questions in the same order; each question is linked to a job-relevant skill or competency identified through job analysis; and applicant responses are evaluated using the same scoring scale and procedure.

Structured interviews use two basic types of questions. *Past-focused* questions require applicants to describe their performance in a job-relevant situation from their past experience. For example, a potential past-focused structured interview question for the competency *interpersonal skills* could be, "Describe a time when you had to work with an uncooperative coworker on a project. What was the situation? What did you do, and why? What was the outcome?" They share the challenges of T&E assessments in general (which are reviewed in the next chapter). Their advantage, in addition to having a high predictive validity, is the opportunity for a live demonstration of performance by the applicant, and a chance to clarify any ambiguities in the questions, albeit in a structured manner.

Present-oriented structured interview questions, on the other hand, ask applicants to demonstrate their current proficiency by describing what they would do in a hypothetical situation.   For example, a potential present-oriented structured interview question for the competency *interpersonal skills* could be, "Describe what you would do if a colleague was dominating the discussion in staff meeting, not leaving others time to talk? What actions would you take and why? What outcome would you expect?" These questions are not T&E measures. Although applicants may draw from past experiences when responding to such questions, present-oriented structured interview questions ask applicants to perform in the present rather than just report on their past. There are a different set of issues to be considered and best practices for present-oriented assessments.

Structured interviews are flexible, easy to develop, and widely used in both private and public sector hiring. MSPB's report, *The Federal Hiring Interview: Unrealized Potential*, reviews best practices for structured interviews.[61] The Office of Personnel

---

[60] Guion, R. (2011). *Assessment, measurement, and prediction for personnel decisions.* Routledge: New York, NY, p. 493.

[61] MSPB (2003). The Federal Selection Interview: Unrealized Potential. MSPB: Washington, DC.

Management also offers training[62] and guidance[63] to assist hiring managers in using structured interviews effectively.

Validity studies have not distinguished between past- and present-oriented structured interview questions because they are often used together in interviews. For this reason it is not possible to report a separate validity coefficient for past-focused structured interviews.

Biodata. Biodata measures bridge the "Likely Useful" and "Very Useful" categories for predicting future job performance (see Table 1). They achieve validities in the vicinity of .35. Some questions on a biodata questionnaire ask applicants to "… recall and report their typical behaviors or experiences in a situation likely to have occurred earlier in their lives."[64] These situations include work experiences, but may also ask about volunteer and relevant non-work experiences that are indicators of abilities relevant to work performance. Applicant behavior in these past situations should be observable, verifiable, and under the person's control rather than just something that happened to them.[65] These biodata questions are similar to past-focused structured interview questions, except that they are administered in questionnaire form with response options without the opportunity to interact with the questioner.

Biodata instruments are not pure T&E assessments because these past-focused questions are not the only kind of question they contain. Some biodata questions address attitudes, motivation, interests, and personality traits.[66] Applicants have mixed reactions to biodata measures, some finding the questions intrusive and inappropriate. Some Federal hiring programs have used biodata measures in the past,[67] but such a decision must be made cautiously. Complexities in development, scoring, and applicant reactions require professional expertise and thorough field testing before deploying a biodata instrument.

[62] See www.opm.gov/services-for-agencies/assessment-evaluation/structured-interview/ and www.hru.gov for structured interview training developed by OPM.

[63] See www.opm.gov/policy-data-oversight/assessment-and-selection/structuredinterviews/ for OPM guidance on the use of structured interviews.

[64] Mumford, M., Whetzel, D., Murphy, S. & Eubanks, D. (2007). Background data. In D. Whetzel and G. Wheaton (Eds.), Applied Measurement: Industrial Psychology in Human Resources Management (pp. 161-180). Lawrence Erlbaum Associates: Mahwah, NJ.

[65] Guion, R. (2011). *Assessment, measurement, and prediction for personnel decisions.* Routledge: New York, NY, p. 512; Mael, F. (1991). A conceptual rationale for the domain and attributes of biodata items. *Personnel Psychology*, 44, 763-792.

[66] Mumford, M. (1999). Construct validity and background data: Issues, abuses, and future directions. *Human Resource Management Review*, 9(2), 117-145.

[67] Gandy, J., Dye, D., & MacLane, C. (1994). Federal government selection: The individual achievement record. In H. Stokes, M. Mumford & W. Owens (Eds.), *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction.* (pp. 275-309). CPP Books: Palo Alto, CA.

Problematic T&E Assessments

Several measures of training and experience are widely accepted because they are easy to use and appear to reflect applicant possession of work-related competencies. Although not an exhaustive list, examples include tallies of years of experience, number of training classes attended, class rank or grade point average, and stand-alone resumes. Despite their intuitive appeal, research has demonstrated that such measures have limited usefulness in predicting job performance. They are reviewed here to ensure awareness of their shortcomings and to highlight potential "red flags" that decision makers should look for when evaluating the quality of any assessment. Where such methods and red flags appear in hiring processes, consideration should be given to replacing or supplementing them with higher-quality measures.

Years of job experience. With an average validity coefficient of .18, years of experience is in the "Possibly Useful" category for predicting future job performance (see Table 1). There is an intuitive appeal to the idea that the length of time an employee has spent doing a certain kind of work can be used as a measure of the employee's skill in doing such work. It can be tempting to translate this directly into a T&E assessment by creating a tally and treating it as a score. It is quick and easy. And, there is some indication that length of experience may predict job success somewhat more accurately when all applicants have less than three years of experience—when all experience falls within the initial steep portion of the on-the-job learning curve. However, after the first two or three of years of work, the value of years of experience has little practical value.[68]

Overall though, the amount of time someone has spent working does not predict future job performance very well. Using year count as an assessment requires making fallacious assumptions such as that more experience automatically produces an increase in competence and performance and that all people develop at more or less the same steady rate. As we will discuss in the next chapter, there are reasons to doubt that such assumptions are true.

Number of training classes attended. There is no validity evidence indicating that the number of training courses attended is useful in predicting future job performance. Mere presence in a course is no guarantee of learning, and therefore awarding points for attending such classes is a poor substitute for evaluating actual learning. Training classes rarely have good end-of-course assessment that might provide evidence of how much learning has occurred.[69] Even when they exist,

---

[68] McDaniel, M., Schmidt, F. & Hunter, J. (1988). A meta-analysis of the validity of methods for rating training and experience in personnel selection. *Personnel Psychology*, 41, 283-314.

[69] Arthur, W., Bennett, W., Edens, P. & Bell, S. (2003). Effectiveness of training in organizations: A meta-analysis of design and evaluation features. *Journal of Applied Psychology*, 88(3), 234-245.

such assessment scores are almost never available to reviewers of job applications.[70] Further, some competencies cannot be effectively developed through training. For such competencies, training classes should especially not be used as a proxy for applicant proficiency.[71]

Overall, MSPB maintains that, "Selection officials should avoid using training and experience measures that credit applicants for competencies on the basis of lists of classes they have attended. Such measures are particularly poor evidence that applicants have acquired less trainable competencies."[72] In doing so, MSPB concurs with OPM's conclusion: "Past behaviors, not past exposure, are the best predictors of future behaviors."[73]

Transcripts and GPA. Grade point average is typically computed using a 0 to 4.0 scale. On the surface, GPA seems like it should predict job performance reasonably well. GPA is easy to obtain, understand, and verify, and it seems to reflect long-term, sustained performance. There is overlap between the skills needed to perform well in school and on the job; the level of motivation and conscientiousness required for school performance seems likely to transfer to job performance; and grades in a major or area of concentration are reasonable indicators of the level of knowledge in that area.[74] Applicants who have received instruction in a particular area require less training on the job. But that relationship only holds for a year or two following the applicant's educational experience.

Yet, despite its intuitive appeal, research places overall GPA predictive validity around .17.[75] Therefore, GPA is in the "Possibly Useful" category for predicting future job performance (see Table 1). A host of factors make GPA an unreliable measure. They include dissimilarity of school environments and work environments; different

---

[70] Even when end-of-course assessments are conducted, they suffer from data quality problems. Analysis of survey responses to the MSPB 2010 indicates that there are many non-training reasons that people attend training that have little to do with learning. Some are unjustified, such as travel to a desired location or avoiding unpleasant work assignments. Others serve a legitimate purpose, such as evaluating the training, preparing to teach it, or networking with other participants. These and other factors affect the scores on end-of-training assessments, undermining their usefulness on the few occasions when they are used.

[71] MSPB (2010). Making the Right Connections: Targeting the Best Competencies for Training. MSPB: Washington, DC.

[72] MSPB (2010). Making the Right Connections: Targeting the Best Competencies for Training. MSPB: Washington, DC, p. 43.

[73] Schmidt, F., Caplan J., Bemis, S., Decuir, R., Dunn, L. & Antone, L. (1979). The behavioral consistency method of unassembled examining (TM-79-21). Washington, DC: U.S. Office of Personnel Management, Personnel Research and Development Center.

[74] Baird, L. (1983). The role of academic ability in high-level accomplishment and general success. College Board Report 82-6. College Entrance Examination Board: New York, NY.

[75] Schmidt, F. & Hunter, J. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262–274.

grading policies; grade inflation; student ability level; self-selection out of difficult classes; and differences in how GPAs are calculated.[76] Further, research has found that the predictive validity is not only low, but seems to have declined over the last two decades.[77] A large meta-analysis of studies relating GPA to a number of work performance measures found little predictive validity for high school and college GPA—although grades in classes taken post college degree (e.g., graduate school) had higher predictive validity.[78] There is some indication that academic achievement is better at predicting training success on the job than work performance. There is also some indication that under some specialized circumstances, GPA can be predictive of job performance. Overall though, the accumulated evidence suggests that GPA is not as useful as many Federal hiring managers believe it to be.[79]

Resumes.   Resumes, "…typically describe the major duties associated with each position in the [individual's] work history, specific coursework, special skills and credentials, and other qualifications."[80] There are some common expectations about what appears on a resume, but much is left up to the applicant.

Resumes are one of the most widely-used pre-employment assessments in private, public, and nonprofit sectors. Potential employers are greatly influenced by resumes, including aspects such as extracurricular activities and visual appeal which have little relevance to future job performance.[81] MSPB found in 2000 that 96 percent of Federal supervisors reported using work history to a great or moderate extent when making hiring decisions.[82] A recent survey of Federal HR practitioners found that more than 90 percent of hiring decisions were based, to some degree, on information from applicants' resumes.[83]

---

[76] Schneider, R. (1984). The rating of experience and training: A review of the literature and recommendations on the use of alternative E & E procedures. IPMAAC Personnel Assessment Monograph.

[77] Roth, P. & Shippmann, J. (1996). Meta-analyzing the relationship between grades and job performance. *Journal of Applied Psychology*, 81(5), 548-556.

[78] Bretz, R. (1989). College grade point average as a predictor of adult success: A meta-analytical review and some additional evidence. *Public Personnel Management*, 18(1), 11-22.

[79] MPS 2000, Question 77A.

[80] McGonigle, T. & Curnow, C. (2007). Measures of training and experience. In Applied Measurement: Industrial Psychology in Human Resources Management. Lawrence Erlbaum Associates: Mahwah, NJ.

[81] Chen, C., Huang, Y. & Lee, M. (2011). Test of a model linking applicant resume information and hiring recommendations. *International Journal of Selection and Assessment*, 19(4), 374-387.

[82] MPS 2000, Question 77A

[83] MSPB Fair and Open Competition Survey (See http://mspb.gov/studies/surveys.htm).

However, resumes have an average predictive validity of only .11,[84] which means they bridge the "Unlikely to be Useful" and "Possibly Useful," categories for predicting future job performance (see Table 1). The validity of resumes depends not only on quality of information provided, but also on how a set of resumes is evaluated and scored. They are often scored using the *holistic method*. The reviewer looks through a stack of resumes and forms a general impression of each applicant. Unsurprisingly, the holistic method is an ineffective way of identifying applicants who perform well on the job. In an unstructured evaluation, such as reviewing a resume without guidelines, people adopt informal strategies that produce poor results. For example, there is a "more is good" effect in which they unconsciously give more credit for resumes that provide a lot of detail—even if much of it is irrelevant.[85]

Resumes can also be scored using a point method which gives credit for various indicators of experience. Point methods tend to focus attention on quantity of experience at the expense of the quality of that experience.[86] OPM notes that resumes have a "…particularly weak ability to predict job performance when scoring method gives credit for factors such as: length and recency of education, academic achievement, and extracurricular activities."[87]

Overall, job applicants are wise to craft their resumes carefully; hiring managers would be wise to give them less weight. It is useful to consider resumes (and application forms, which are basically standardized resumes), not as poor T&E assessments, but as incomplete T&E assessments. While the unverified information in a stack of resumes does not predict future performance, it does contribute to such prediction when combined with a reference check that confirms or contradicts the information.

Indicators of a low quality T&E assessment. The aforementioned examples of problematic T&E assessments were included to highlight points of caution in T&E assessments. This was not an exhaustive list of problematic T&E assessments; decision makers will need to exercise wise judgment in determining the appropriateness of a given T&E assessment for their particular hiring needs or situation. They are responsible for identifying the disadvantages of using any particular T&E assessment, and for weighing such information against any perceived advantages. As discussed above, red flags which warrant extra consideration include:

---

[84] Schmidt, F. & Hunter, J. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262–274.

[85] Highhouse, S. (1997). Understanding and improving job-finalist choice: The relevance of behavioral decision research. *Human Resource Management Review*, 74, 449-470.

[86] McGonigle, T. & Curnow, C. (2007). Measures of training and experience. In Applied Measurement: Industrial Psychology in Human Resources Management. Lawrence Erlbaum Associates: Mahwah, NJ.

[87] OPM, *Assessment Decision Tool*.

- *Low or Unreported Validity.* Hiring officials and HR professionals should be cautious about any assessment which has not been subjected to a validation procedure to determine how well it predicts job performance. Too often measures of training and experience may appear rational and valid, but they have not been put to the test. The best approach is to avoid the use of any such measures until their validity is established. Whether or not they are eventually shown to produce valid results, their use can be difficult to defend in the face of a legal challenge without validity evidence in hand.[88]

- *Exclusive Focus on the Passage of Time.* Practitioners should particularly avoid any methods which create a score based solely on the passage of time. There is no assurance that time or attendance indicates learning. Consequently, such measures are unlikely to be good predictors of job performance. Measures based on length of experience are pervasive in minimum qualifications used as the first hurdle in much of Federal hiring. When such measures must be part of a hiring process, it is important to check the results with a more accurate measure later in the assessment process.

- *Exclusive Focus on the Amount of T&E.* Using measures which value the *amount* of training and experience with little attention to whether it resulted in any improved ability to perform work can result in hiring less qualified applicants than those who may have spent less time in class but have stronger abilities. Similarly, promising entry-level applicants can be overlooked because their potential to learn new skills quickly is not considered or because they learned from life experiences that a narrowly-targeted training-focused T&E measure may not recognize.

Summary

This chapter reviewed both useful and problematic T&E assessments commonly used in Federal hiring and promotion decisions. Some, such as occupational questionnaires and accomplishment records, can contribute to hiring decisions. Similarly, structured interviews and biodata, which can include T&E information,

---

[88] *See Albemarle Paper Co. v. Moody*, 422 U.S. 405 (1975). In *Albemarle*, the employer used two different tests to assess employees for promotion, but did not establish at the time of use what connection, if any, there was between the tests and the jobs at issue. *Id.*, at 410-11. "[O]n the eve of trial, Albemarle engaged an industrial psychologist to study the 'job relatedness' of its testing program." *Id.* at 411. The Supreme Court noted that "It cannot escape notice that Albemarle's study was conducted by plant officials, without neutral, on-the-scene oversight, at a time when this litigation was about to come to trial. Studies so closely controlled by an interested party in litigation must be examined with great care." *Id.* at n. 32. Ultimately, Albemarle lost the case because it did not adequately prove the tests it used were related to the jobs for which they were used. *Id.* at 435-36.

are also useful in gauging an applicant's likelihood of success in a job. However other T&E assessments, like years of work experience or training class attendance,provide little reliable information about an applicant's abilities and should be replaced or supplemented by higher-validity measures.

While each method has its own profile of strengths and weaknesses, there are some challenges common to all T&E assessments. The next chapter reviews several of these challenges.

Chapter 3 - What Are the Common Challenges?

The T&E assessment methods reviewed in the last chapter differ in a number of ways, including their overall quality, what they require of applicants, how they are scored, and the level of resources necessary to develop and administer them. But as measures that target training and experience, they also face common challenges. Such challenges include focusing on the past; applicant difficulty evaluating their own abilities; and applicant dishonesty. It is useful to identify these challenges and highlight how each one can impact the ability of a specific T&E assessment to predict the future job performance of applicants. Decision makers should consider this information when determining whether or not a particular T&E assessment is appropriate for their hiring needs and situation.

Focus on the Past

There are some commonsense reasons that a T&E assessment's focus on the past seems advantageous. T&E measures are grounded firmly in the principle of *behavioral consistency*.[89] "A candidate who performed well on a job in the past is likely to perform well on a similar job in the future; one who has behaved responsibly in the past is likely to be responsible in the future."[90]

There is a strong sense that the past already exists, cannot be changed, and is therefore easy to verify. Further, many T&E assessment methods require minimal development effort and can be put into place quickly.

Applicants also seem to agree that T&E measures are reasonable, or have good face validity. As a result, hiring decisions based on them are challenged much less often than other types of assessments.

Despite this allure, there are three kinds of challenges with trying to predict future job behavior from what we are told about the past: (1) what an applicant reports

---

[89] McGonigle, T. & Curnow, C. (2007). Measures of training and experience. In D. Whetzel and G. Wheaton (Eds.), Applied Measurement: Industrial Psychology in Human Resources Management (pp. 161-180) Lawrence Erlbaum Associates: Mahwah, NJ.

[90] Guion, R. (2011). *Assessment, measurement, and prediction for personnel decisions*. Routledge: New York, NY, p. 508. See also Schmidt, F., Caplan, J., Bemis, S., Decuir, R., Dunn, L., & Antone, L. (1979). The behavioral consistency method of unassembled examining. Washington, DC: U.S. Office of Personnel Management, Personnel Resources and Development Center.

is not necessarily what happened; (2) applicants do not necessarily learn effectively from past events; and (3) training and experiences in the past may not be applicable to now, or to the future.

Challenge 1: What an applicant reports is not always what happened. The past may be unchangeable, but our access to it is indirect and imperfect. Our assumptions about what happened during training or a work experience may be incorrect. Simple attendance at training is no guarantee that the attendee paid attention or learned effectively. This may be true for reasons unrelated to the attendee, including poorly designed or delivered training, training not targeted at the right level (e.g., introductory courses when advanced courses are needed), and dated or inaccurate training content. Even when a person is prepared, motivated, and engaged, learning does not occur at the same rate over time. It typically follows a curve of diminishing returns where a great deal is learned initially and the amount learned decreases over time, approaching some "mastery maximum" where most possible learning has occurred. This curve will have a different shape for different skills, environments, and a number of other factors.[91]

Performance-improving experience also may not be obtained on the job. Much of an employee's work may have been performed by others, performed poorly, or performed under very different conditions or levels of assistance than implied. Even if performance level is not an issue, some abilities may deteriorate if they are not used. It is also not uncommon for a job title to imperfectly reflect the duties actually performed. The unused skills implied by the job title may have once been possessed by an employee, but may have atrophied through disuse during the "years of experience."[92]

Time passage is a particularly poor measure of whether an individual possesses the kind of abilities that develop or emerge only in the face of unusual, highly-demanding situations.[93] For example, firefighters who have been through a southern Californian brushfire season have a very different year of experience than they would in a rainy eastern city with only an occasional localized fire.

Further, applicant descriptions of their past experience often conceal, or at least deemphasize, gaps in their periods of employment. While omitted events could mean something irrelevant to the job, like family medical leave, they could also

---

[91] Goldstein, I. and Ford, K. (2004). Training in organizations: Needs assessment, development and evaluation. Wadsworth: Belmont, CA.

[92] Jacobs, R., Hoffman, D. & Kriska, S. (1990). Performance and seniority. *Human Performance*, 3, 107-201.

[93] Elder, G. & Clipp, E. (1989). Combat experiences and emotional health: Impairment and resilience in later life. *Journal of Personality*, 57, 311-342.

include something job-relevant, such as a period of incarceration. A history gap often indicates a period of unemployment, a time during which no kind of experience is being gained. It can be difficult to establish how many years really were spent gaining experience.

Challenge 2: Applicants do not all learn effectively from past training and experiences. There are individual differences in how people acquire, retain and recall information. People come to both training and work experiences with different levels of various abilities. This affects how quickly they learn and, in some cases, what they are able to learn.[94] It has been suggested that there are three primary factors which influence how much two different people might learn from the same experience in the same length of time. These three factors are: "… individual differences, such as cognitive ability, conscientiousness, openness to experience, and extraversion; situational characteristics such as opportunities to perform tasks; and motivation to pursue opportunities."[95] Individuals learn and develop at different rates, making it difficult to estimate how much a given applicant may have improved based on work experiences—and how quickly such skills may atrophy when they are not exercised for a time.[96] This is consistent with previous MSPB studies highlighting that people differ in their preparedness and ability, resulting in differential benefits from the same training.[97]

Also, length of time at a job might not always be a good thing. In some cases, the length of time spent in a job may be negatively related to ability level, if it reflects an inability to obtain advancement or another position.[98] "The idea that anyone can become expert given enough time is a myth; those who take a long time to learn a task generally do not reach the level of proficiency after training reached by those who learn it more quickly."[99] The risk that time on the job reflects lower, rather than higher, proficiency is greater if the experience was gained in an organization that does not effectively address performance problems, for reasons that may include indecision or reluctance to invest the organizational effort needed to remediate or remove a poor performer.

---

[94] MSPB (2010). Making the Right Connections: Targeting the Best Competencies for Training. MSPB: Washington, DC.

[95] McGonigle, T. & Curnow, C. (2003). Current theory and practice on the measurement of experience. Paper presented at the Mid-Atlantic Personnel Assessment Consortium. Rehobeth Beach, DE.

[96] P. L. Ackerman. (1988). Determinants of individual differences during skill acquisition. *Journal of Experimental Psychology*, 117, 288-318; and Thoresen, C., Bradley, J., Bliese, P. & Thoresen, J. (2004). The big five personality traits and individual job performance growth trajectories in maintenance and transitional job stages. *Journal of Applied Psychology*, 89, 835-853.

[97] MSPB (2010). Making the Right Connections: Targeting the Best Competencies for Training. MSPB: Washington, DC.

[98] Jacobs, R., Hoffman, D. & Kriska, S. (1990). Performance and seniority. *Human Performance*, 36, 107-201.

[99] Guion, R. (2011). *Assessment, measurement, and prediction for personnel decisions.* Routledge: New York, NY.

Further, as we will discuss in greater depth below, people have various kinds of memory problems recalling past experiences. And people suffer from these problems to different degrees and have differing degrees of success using strategies to mitigate them.[100]

Challenge 3: Applicants' past training and experiences may no longer be relevant. Research suggests that the value of training and experience gradually decreases with the passage of time as forgetting occurs, skills become rusty with disuse, and technology and best practices evolve in the workplace.[101]

Jobs also differ in the degree to which they require skills and knowledge from the past. "While experience may be more important where success depends heavily upon procedural or institutional knowledge, it may be less important where the needed KSAs are rapidly shifting, particularly when the employee does not practice continual learning in an effort to stay current in his or her field."[102] Jobs and professional disciplines change at different rates over time. Ironically, the very climate of speed and change that encourages managers to adopt T&E assessments also contributes to the greater speed at which work skills become obsolete.[103]

In summary, these three challenges are inherent to focusing on the past for information about job-related performance and abilities. Below, we discuss an additional set of challenges that emerge when we ask candidates to evaluate their experience.

Unreliable Self-Assessment

It is generally accepted that self-ratings tend to overestimate experience, ability, and performance, and do not distinguish well between qualified and unqualified applicants.[104] Applicants can have problems recognizing, supplying, and in some cases, evaluating information about their training and experience. These issues are based on shortcomings in human abilities to perform the activities required by many T&E assessments. Below, we will explore three main challenges in applicant self-assessment: (1) applicants may have trouble recalling or reporting the most

---

[100] Schacter, D (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist* 54 (3), 182–203.

[101] McDaniel, M., Schmidt, F. & Hunter, J. (1988). A meta-analysis of the validity of methods for rating training and experience in personnel selection. *Personnel Psychology*, 41, 283-314.

[102] MSPB (2013). Preserving the integrity of the Federal Merit Systems: Understanding and addressing perceptions of favoritism. MSPB: Washington, DC.

[103] Jansen, P. (1997). Assessment in a technological world. In N. Anderson and P. Herriot (Eds.), International Handbook of Selection and Assessment (pp. 125-145). Chichester, UK: John Wiley.

[104] van Rijn, P. (1980). Self-assessment for personnel examining: An overview. (OPRD Report 80-14). OPM.

appropriate or applicable experiences; (2) applicants may not have an accurate perspective on the range of their proficiencies; and (3) applicants may not accurately evaluate their proficiencies. Fortunately, many of these issues can be corrected.[105]

Challenge 1: Applicants have problems recalling past experiences. T&E assessment questions often ask applicants to recall and evaluate the past, reflecting both the necessity and the assumption that applicants are the best sources for this information. Yet, people experience difficulty and make systematic errors when trying to remember their past experiences.[106] Research indicates that recall of past events is not only incomplete, but may be influenced by people's mistaken beliefs that their past behavior is consistent with their current behavior, and with their own conceptions of ideal behavior.[107]

Further, applicants may not always report the most appropriate or applicable experiences. For example, some T&E assessments impose brevity constraints that require applicants to report only a subset of their achievements. The achievements described are not necessarily the ones that show the upper range of applicants' abilities.[108] Applicants may inappropriately focus on the accomplishments that are the most recent; those that stand out because they were early in their career; or those that were enjoyable, interesting, or otherwise memorable. Unfortunately, such achievements might not best represent applicants' full range of capabilities or best match the kinds of experiences that a hiring manager is looking for.

Also, such decisions about which experiences to report can also be affected by the number of options people must choose from.[109] While applicants with novice experience in an area may be able to easily recall and report the few opportunities they have had in a particular area, the applicant with expert experience has the far more difficult task of selecting from a much larger number of accomplishments; it may not be easy to come up with one that shows the expert applicant in the best possible light for the position.

Challenge 2: Applicants lack perspective on their level of proficiency. Readers may be familiar with the phrases, "I know enough to know that I know nothing,"

[105] Podsakoff, P., MacKenzie, S. & Podsakoff, N. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879-903.

[106] Schacter, D. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist* 54 (3), 182–203.

[107] Pearson, R., Ross, M. & Dawes, R. (1992). Personal recall and the limits of retrospective questions in surveys. In J. M. Tanaur (Ed.), *Questions about questions: Inquiries into the cognitive bases of surveys* (pp. 65-94). Russell Sage Foundation: New York, NY.

[108] Kahneman, D. (2012). Thinking, fast and slow. Farrar, Straus & Giroux: New York, NY.

[109] Guion, R. (2011). *Assessment, measurement, and prediction for personnel decisions.* Routledge: New York, NY.

or "I know just enough to be dangerous." These derive from the reality that people have difficulty understanding how much they know about a given area. Many errors that people make in self-ratings occur because they lack perspective about their own abilities. People differ in the degree to which they are reflective and able to achieve accurate self-knowledge.[110] Further, low and high performers differ in their awareness of the frontier of knowledge in a particular area. Low or novice performers are unfamiliar with the nature of advanced performance—this is part of what makes them low performers—and they may incorrectly believe they are close to the top of the performance dimension. They know very little about what they do not know and they often cannot fully describe what is required on the job.[111] Yet, even top performers have trouble comparing themselves to others, and tend to *underestimate* their own abilities.[112] High performers are aware of the frontier of the unknown and the more complex problems they might face. This makes them less likely to endorse that their level of knowledge or experience is at the top of the continuum; they realize the truly advanced nature of expert level.

Challenge 3: Applicants have problems self-evaluating. Coupled with often having poor perspective on the range of their abilities, applicants also have difficulties rating their own abilities. Sources of error in self-ratings include differences between applicants in memory as they attempt to recall their previous experiences, and differences in the degree to which individuals are reflective and have accurate self-knowledge about their work-related performance.[113] Further, people differ in their abilities to self-rate and in other characteristics that affect the quality of these ratings. Factors identified by research include people's beliefs about their abilities to accomplish tasks,[114] conscientiousness (as measured by a personality inventory),[115] general intelligence, and their cognitive complexity when rating on multiple dimensions.[116]

---

[110] Graham, K., McDaniel, M. & Snell, A. (2002). Biodata validity decay and score inflation with faking: Do item attributes explain variance across items? *Journal of Business and Psychology*, 16(4), 573-592.

[111] IAG-CTEA, (2009).

[112] Dunning, D., Heath, C. & Suls, J. (2004). Flawed self-assessment: implications for health, education, and the workplace, *Psychological Science in the Public Interest*, 5(3), 69-106; and Thornton, G. (1980). Psychometric properties of self-appraisals of job performance. *Personnel Psychology*, 33(2), 363-271.

[113] Graham, K., McDaniel, M. & Snell, A. (2002). Biodata validity decay and score inflation with faking: Do item attributes explain variance across items? *Journal of Business and Psychology*, 16(4), 573-592.

[114] Bernardin, H. J. and Villanova, P. (2005). Research streams in rater self-efficacy. *Group and Organization Management*, 30, 61-88.

[115] Tziner, A., Murphy, K., and Cleveland, J. (2005). Contextual and rater factors effecting rating behavior. *Group and Organization Management*, 30, 89-97.

[116] Guion, R. (2011). *Assessment, measurement, and prediction for personnel decisions.* Routledge: New York, NY.

Additionally, applicants are unlikely to have received training in how to use rating scales to effectively calibrate their proficiencies. Unfortunately, untrained raters tend to form inaccurate personal impressions about rating scales and how to apply them. Such inaccuracies result in distortions in how applicants appraise and rate their proficiencies. For example:

- Central tendency bias occurs when ratings cluster closely around the middle of the scale;

- Leniency or severity bias occurs when raters use either the top or the bottom of the scale, making few distinctions; and

- Halo bias occurs when a general impression is used to self-rate across a set of specific tasks.

While there are statistical methods to detect and correct for these biases, such methods require data from other raters for comparison. In self-rating, this is not possible as rating data only comes from one source: the applicant.

Challenges in appropriately recalling relevant performance information, maintaining proper perspective, and correctly rating abilities are faced by applicants who attempt to report and evaluate their experience fairly, accurately, and completely. Unfortunately, not all applicants have such good intentions, leading to another area of challenges with T&E assessments: deliberately false or misleading assertions by candidates about their training and experience.

Applicant Dishonesty

Some applicants deliberately fabricate or exaggerate their past training, experience, and accomplishments. One review of studies that examined applicant dishonesty estimated that 30 to 70 percent of job applications contain untrue statements or significant exaggerations.[117] Another study found that between 1 and 78 percent of occupational questionnaire scores for eligible applicants across six Federal job openings were not supported by an applicant's documentation. While a variety of factors likely contributed to these results, exaggeration and dishonesty likely contributed as well.[118]

Some assessment developers discuss an *inflation bias* that occurs when applicants portray themselves in the best possible light, perhaps without realizing they are

---

[117] Wood, J. L., Schmidtke, J. M. & Decker, D. L. (2007). Lying on job applications: The effects of job relevance, commission, and human resource management experience. *Journal of Business and Psychology*, 22, 1-9.

[118] Barton, M. (2010). Automated T&E questionnaires: Practical outcomes and development considerations. SIOP.

doing so.[119] This is part of a general truthfulness problem researchers call *social desirability* in survey research and *faking* in personality assessment. It is a significant threat to accurate T&E assessment.

Reference checking is a widely-used attempt to reduce the impact of applicant dishonesty. MSPB has studied its use in Federal hiring and has found that it can be accomplished using methods that are both legally defensible and effective.[120] Yet, reference checking must be structured and thorough to be effective. Unstructured reference checking, wherein a hiring manager asks unplanned and inconsistent questions across candidates, is not wise and unlikely to provide useful information in a balanced, transparent, or fair manner. Seeking only confirmatory information is also unwise; irrespective of the impressions a hiring manager has about an applicant's qualifications, hiring managers should be open to all information during reference checking, both good and bad. Indeed, hiring managers need to avoid falling into the trap of looking to confirm their preconceptions, instead of testing them or disproving them.[121]

Applicant dishonesty, abetted by ineffective reference checking, remains a challenge to accurate T&E assessment.[122]

Summary

This chapter examined challenges encountered when evaluating past training and experience. Some stem from the nature of the past. Others reflect problems in *how* the past is probed. The next chapter reviews strategies that can be used to either improve the effectiveness of agency T&E assessments, or identify alternative methods.

---

[119] Barton, M. (2010). Automated T&E questionnaires: Practical outcomes and development considerations. SIOP.

[120] MSPB (2005). Reference Checking in Federal Hiring: Making the Call. MSPB: Washington, DC.

[121] Johnson-Laird, P. (2006). How we reason. Oxford University Press: Oxford, UK.

[122] Some agencies are proactive about evaluating and improving their reference checking practices. See *Reference Checking in the Department of Justice* (2013) available at www.justice.gov/oig/reports/2013/e1302.pdf .

Chapter 4 - How Can We Raise the Bar?

The last chapter highlighted challenges faced when using training and experience assessments. Most of these are not fatal flaws. Significant improvements can be made in how well T&E assessments predict future job performance if the right techniques are used to counter their weaknesses.[123]

In this chapter we review several strategies for improving T&E assessments. These strategies are:

- Improve accuracy of T&E assessments;

- Increase verification of applicant information;

- Use present-oriented assessments; and

- Monitor the latest assessment information and trends.

We discuss these strategies at a general level, but caution readers that such strategies must be adapted to the specifics of each type of assessment. Further, although these techniques have the potential to mitigate several weaknesses in T&E assessments, hiring managers are still responsible for exercising wise judgment to identify the most appropriate assessment for their individual hiring needs and situations. This includes weighing the disadvantages and any possible strategies to improve any assessment, against its advantages. This analysis is not easy, and hiring managers might consider guidance from consultants or HR specialists who specialize in assessment development and use.

Improvement Strategy 1: Make T&E Assessments More Accurate

A number of techniques can be applied to increase the accuracy of T&E assessments. They include grounding T&Es in job analysis and improving T&E questions, rating scales, scoring, and applicant self-assessment. Most involve improving how applicants recall and evaluate their own training and experience.

---

[123] See, for example, Quinones, M., Ford, K. & Teachout, M. (1995). The relationship between work experience and job performance: A conceptual and meta-analytic review. *Personnel Psychology*, 48, 887-910.

Ground in job analysis. Among other criteria, useful T&E measures focus on job-relevant attributes. A current job analysis is necessary to establish job-relevant attributes, such as knowledge, skills, abilities, and competencies. [124] MSPB has recommended: "…the Government should be developing assessments that better measure the quality of the skills actually developed through that training and experience. This can be accomplished to a certain extent through T&E assessments, but only if agencies improve the way they do job analyses and crediting plans. In particular, the Government needs to use more rigorous procedures to determine performance elements that separate high performers from low performers…"[125] Agencies that have used T&E measures extensively emphasize the importance of using questions grounded in job analysis.[126]

Ideally, a current, detailed, and well-documented job analysis will either already be available, or the resources will be available to conduct one. If job analysis documentation is not available, it is important to consult position descriptions, vacancy announcements, performance standards, and training materials to determine how and how *well* the job should be done. Subject-matter experts (SMEs), or those who are familiar with the job as it is currently performed, should also be consulted. The best SMEs are those who are exceptional performers. They may be the *go-to* experts for difficult problems; they may be supervisors or instructors of other employees. In any case, SMEs can be an invaluable resource in generating—or vetting—job analysis material

Improve T&E questions. An assessment can only be as good as the quality of the questions it asks. Some quality standards for questions are obvious and apply to all assessments in general. For example, assessment questions should be clear to both experienced and inexperienced applicants, and they should be as concise as possible. If an applicant cannot understand what a question is asking, it will be very difficult for him or her to provide an accurate or appropriate response, even with the best of intentions. Other question quality standards are more specific to the kinds of questions that appear only in T&E assessments. We provide a sample of such T&E specific standards below. Although these standards can be applied to T&E methods in general, they are most pertinent to occupational questionnaires and accomplishment records.

---

[124] OPM's Assessment Decision Guide defines job analysis as: "A systematic examination of the tasks performed in a job and the competencies required to perform them." See also Prien, E., Goodstein, L., Goodstein, J. & Gamble, L. (2009). A practical guide to job analysis. Wiley: Hoboken, NJ.

[125] MSPB (2006). Reforming Federal Hiring: Beyond Faster and Cheaper. MSPB: Washington, DC.

[126] Agarwal, S., Busciglio, H., Leaman, J., Simmons-Collins, T., Curtin, P. & Thompson, S. (2010). Understanding the practical outcomes of self-report, task-based assessment questionnaires for assessing applicants for U.S. Customs and Border Protection vacancies. SIOP, 2010.

- **Stimulate recall.** When possible, questions and instructions should be written in a way that helps stimulate accurate recall of past events. One approach guides applicants through consideration of the length of their experience with a task and the frequency of task performance. This not only helps scoring, but stimulates the applicant's recall of the task and associated context.

- **Distinguish between levels of proficiency.** Items should be written in a manner that allows for clear distinctions between acceptable and unacceptable levels of proficiency. Tasks that either everyone can do or that no one can do provide no useful information and waste the applicant's time. For example, a question that nearly every applicant gets a high score on does not help the agency hone in on the most qualified candidates. In contrast, a question that virtually no candidates can answer is too demanding, and may cause high-quality candidates to self-eliminate out of the application process on the assumption that the employer is seeking—and expects to find—an expert-level candidate.

- **Encourage applicants to review their job qualifications.** To facilitate accuracy in rating their T&E, applicants should review their performance appraisals and similar documents before a T&E assessment to walk through their work history and prepare to answer questions about it.

Developing questions that elicit useful, job-related information from applicants is a specialized skill and is unlikely to be done well without a certain amount of training and experience.[127] As mentioned above, hiring managers should seek guidance from consultants or HR specialists who have specialized knowledge in assessment development.

Improve T&E rating scales. In questionnaire-based T&E assessments, applicants are typically asked to use a rating scale to self-evaluate their abilities or proficiencies. As discussed, the task of accurately recalling and appraising one's past experiences or overall expertise in an area is a challenging process. This challenge is compounded when an individual is given a poorly designed rating scale that does not appropriately capture or convey the full range of proficiency in a given area. A poor rating scale is confusing or difficult to use; a good one is easy to understand, grounded in job analysis, and effectively represents the continuum of proficiency in a given

---

[127] OPM (2011). A training guide for developing assessment questionnaires in USA Staffing. OPM: Washington, DC.

area, while distinguishing between milestone levels of proficiency in that area. Detailed guidance on the proper design of rating scales for T&E assessments is beyond the scope of this report. Hiring managers are encouraged to seek guidance from consultants, HR specialists, or others who have specialized training in scale development. However, we provide a few points to consider when developing T&E assessment scales.

- **Specificity of Focus.** When possible, T&E assessment scales should ask applicants to self-rate their past on specific tasks rather than their abilities in general. The "mental algebra" of estimating how several abstract abilities contribute to performance is very difficult for applicants. Rating specific task performance is an easier and more natural undertaking.

- **Number of Response options.** T&E rating scales generally provide applicants with 2, 3, or 5 response options for each question. Five-point scales capture more information than 2 response options, such as "Yes/No," and are preferred. The only exception is when a T&E question asks about a binary topic, such as possession of a license or certificate necessary to perform the job (e.g., license to practice law or board certified in medicine).[128]

- **Scale Labels.** For scales with three or more points, descriptive anchors on the ends, such as "novice level" and "expert level" make the scale easier for applicants to use and less prone to errors.

- **Scale Point Behavioral Examples.** Greater accuracy and applicant understanding of the self-assessment task can be achieved by using behaviorally-anchored rating scales (BARS).[129] The scale includes descriptions of the behavior that should be expected at different levels of task performance, which are derived from a job analysis or SME opinion. Applicants use the scale by comparing their achievements to the example behaviors, and selecting the behavior (and corresponding scale point) that best mirrors their achievement. BARS helps give applicants a common frame of reference to use

---

[128] Agarwal, S., Busciglio, H., Leaman, J., Simmons-Collins, T., Curtin, P. & Thompson, S. (2010). Understanding the practical outcomes of self-report, task-based assessment questionnaires for assessing applicants for U.S. Customs and Border Protection vacancies. SIOP, 2010.

[129] Smith, P. & Kendall, P. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology*, 47, 149-155.

when evaluating their experiences, which can help avoid problems with different applicant perspectives on how high, low, and middle-range achievement should be defined.

- Response Scale Reusability. Scales can be generic, with the same response options applied to all questions, tasks, or abilities on a T&E assessment, or they may be customized for each question, task, or ability. Both styles have been used successfully. See Figure 1 and Figure 2 for examples of generic and customized response options, respectively. Note how the response options in the Generic scale (Figure 1) could be applied to many abilities or competencies, while the response options for the customized scale (Figure 2) are specific to the competency "Oral Communication."

Figure 1. Sample Item with Generic Scale

Instructions: Select the statement that best describes your training and experience in [insert task].
A. I have not had education, training, or experience in performing this task.
B. I have had education or training on this task but have not yet performed it on the job.
C. I have performed this task on the job. My work on this task was monitored by a supervisor or senior employee to ensure compliance with proper procedures.
D. I have performed this task as a regular part of a job. I have performed it independently and normally without review by a supervisor or senior employee.
E. I have supervised performance of this task or I am normally the person who is consulted by other workers to assist them in doing this task because of my expertise.

Figure 2. Sample Item with Customized Scale

Instructions: Indicate the extent to which you have communicated orally with various levels of employees to obtain and provide information.
A. I have had no experience in performing this task.
B. I have communicated orally with others to obtain or verify information or to provide routine information.
C. I have communicated orally with supervisors, managers, or office personnel to notify them of decisions, problems, or further actions needed, or to explain the organization's programs or services.
D. I have given short oral presentations at departmental/organizational briefings and meetings to convey information on program activities or to describe the impact of new organizational policies on operational responsibilities.
E. I have led briefings or taught courses on highly technical or complex material to audiences such as high-level managers, attorneys, or executives.

Improve applicant self-assessment proficiency. When an applicant is asked to rate his or her experiences or qualifications for a particular job, he or she is acting as a rater. There is an abundance of research on rating processes which has insights for understanding (and improving) how applicants rate their abilities. For example, some self-rating difficulties stem from how well applicants understand the rating scales.[130] Investigations of untrained raters and of the errors they make during training reveal that many use their own personal beliefs of how the rating scales should work when making ratings. Although their individual rating behavior can be quite consistent—they are consistently *inaccurate* with respect to how the scales were designed to be used.[131] Training applicants on how to interpret and use T&E rating scales could help them more accurately rate their job qualifications, thereby making T&E assessments which rely on such ratings more useful. Indeed, the quality of applicant self-ratings will improve when scales and self-rating procedures are more clearly explained to applicants before self-rating occurs.[132]

The most effective training method is *frame of reference training*, which gives raters guided practice using the scales they will be using on the assessment. Having learned

[130] Primoff, E. (1980). The use of self-assessments in examining. *Personnel Psychology*, 33, 283-290.

[131] Uggerslev, K. & Sulsky, L. (2008). Using frame-of-reference training to understand the implications of rater idiosyncrasy for rating accuracy. *Journal of Applied Psychology*, 93(3), 711-719.

[132] Bernardin, H. & Buckley, M. (1981). Strategies in rater training. *Academy of Management Review*, 6, 205-212; van Rijn, P. (1980). Self-assessment for personnel examining: An overview. (OPRD Report 80-14). OPM.

the rationale for the rating task, raters use this perspective in their rating tasks.[133] This experience can be used to offer brief self-rating training to applicants prior to beginning a T&E assessment.[134] The key is that applicants receive instruction in how to use the rating scales, so that they will be in a better position to apply them appropriately to their own experiences and abilities.

There are various ways to deliver self-rating training to applicants. Instruction in self-rating could be included in a mail or email message sent to applicants before they engage with the T&E assessment. Some training can also be included in the assessment session, although there are obvious time limits for this. Another option is to integrate web-based training into USAJOBS. Effective web-based training has been developed for raters in other contexts,[135] and would likely be helpful for applicants too. For example, OPM could develop several modules of such training for different forms of assessments and make them available for agencies to use. For its centrally administered assessments, OPM might even require applicants to complete a training module before taking the assessment. Many agencies already do something like this by including links to the assessment questions applicants will encounter. Although job announcements may be longer than applicants can bear,[136] this is another possible place for training. Regardless of method, however, it is critical that all applicants are provided equal opportunity to utilize such training.

Improve T&E scoring. There are a variety of strategies for scoring T&E assessments, some notably better than others. A comprehensive discussion of the complexities of developing a scoring strategy is beyond the scope of this report. However, we emphasize that a key element in any good scoring strategy is that scores reliably and accurately distinguish between levels of applicant quality. An MSPB review of Federal hiring recommended that, "…the Government needs to use more rigorous procedures to determine performance elements that separate high performers from low performers and to develop valid scoring techniques for measuring applicants against those dimensions."[137]

---

[133] Bernardin, H. & Buckley, M. (1981). Strategies in rater training. *Academy of Management Review*, 6, 205-212; Gorman, C. & Rentsch, J. (2009). Evaluating frame-of-reference rater training effectiveness using performance schema accuracy. *Journal of Applied Psychology*, 94(5), 1336-1344.

[134] Lievens, F. & Sanchez, J. (2007). Can training improve the quality of inferences made by raters in competency modeling? A quasi-experiment. *Journal of Applied Psychology*, 92(3), 812-819. For a discussion of training for raters for assessment centers, see Schleicher, D., Day, D., Mayes, B. & Riggio, R. (2002). A new frame for frame-of-reference training: Enhancing the construct validity of assessment centers. *Journal of Applied Psychology*, 87, 735–746.

[135] Aguinis, H., Mazurkiewicz, M. & Heggestad, E. (2009). Using web-based frame-of-reference training to decrease biases in personality-based job analysis: An experimental field study. *Personnel Psychology*, 62, 405-438.

[136] MSPB. (2003). Help Wanted: A review of Federal vacancy announcements. MSPB: Washington, DC.

[137] MSPB (2006). Reforming Federal Hiring: Beyond Faster and Cheaper. MSPB: Washington, DC, p. 21.

Overall, we recommend that hiring mangers seek and follow the advice of an assessment development specialist when deciding how to score a T&E assessment or an assessment of any type. Not doing so risks poor predictive validity for the assessment.

Improvement Strategy 2: Increase Verification of Applicant Responses

Verifying an applicant's T&E information is much more likely to produce an accurate picture of the applicant's abilities than relying on just the applicant's perspective alone. MSPB has previously discussed the importance of verifying information obtained early in the assessment process.[138] Ideally, all information from every candidate would be subjected to verification. Yet, for practical reasons, it is not often feasible to cross check every bit of information that an applicant provides. As such, agencies need to rely on one or more strategies to corroborate a cross-section of applicants or applicant information. There are a range of strategies to choose from. For example, OPM has developed verification strategies for several types of assessments, most notably occupational questionnaires and accomplishment records.[139] Some strategies are *agency-driven*, in which the agency takes responsibility for checking applicant supplied information. Such agency-driven strategies include:

- Verifying information of a random sample of applicants, reducing the resources required while giving each applicant the same chance of investigation;

- Verifying information from misplaced candidates who score highly on other assessments and are candidates for referral or selection;

- Verifying selective content, such as what is critical to the job, what is most often falsified or exaggerated, or what does not seem consistent with other information;

- Verifying information by inserting bogus (i.e., not real) job tasks and looking for applicants who say they have performed such tasks;

- Verifying through additional assessments that seek to corroborate the legitimacy of an applicant's assertions; and

---

[138] MSPB. (2004). Identifying Talent through Technology: Automated Hiring Systems in Federal Agencies. Washington, DC.

[139] Sharpe, P. & Sorensen, K. (2011). Assessment of training and experience. IPAC 2011, Washington, DC. See also Barton, M., Bisges, J. & Holloway-Lundy, A. (2010). Practical outcomes associated with use of automated questionnaires at the U.S. Office of Personnel Management. SIOP 2010.

- Verifying information by using reference checks to consult with people who are familiar with applicants' behavior in past situations.

Agencies can also use verification strategies that are more *applicant-driven*. Such strategies focus on encouraging honest and accurate applicant reporting, placing the onus on applicants to verify (and certify) that what they have written is true. Examples of applicant-driven strategies, include:

- Provide notice of verification or actual warnings about the consequences of misrepresenting or falsifying application information;

- Requiring documentation from applicants that supports their claims or proficiencies; and

- Requiring elaboration from applicants on their experiences to justify their proficiency ratings.

The following subsections highlight several agency-driven and applicant-driven strategies. We provide such strategies for illustrative purposes only; hiring managers are responsible for exercising wise judgment in deciding which verification strategy is most appropriate for their situation.

Notice of verification. One applicant-driven strategy is to encourage applicants to provide more accurate information, and to exaggerate and falsify less, by telling them that their information will be verified. Warnings about verification and negative consequences of falsification have been shown to decrease falsification even when no actual verification occurs.[140] Negative consequences in warnings may include removal from consideration, removal from position after hire, and in some cases prosecution.[141] The best results can be achieved by placing the warning at the beginning of the T&E assessment and again at the end with a reminder that applicants can go back and change their answers.[142] It can also be placed in the vacancy announcement.

---

[140] Dwight, S. & Donovan, J, (2003). Do warnings not to fake reduce faking? *Human Performance*, 16(1), 1-23. See also Lautenschlager, G. J. (1994). Accuracy and faking of background data. In G. Stokes, M. Mumford & W. Owens (Eds.), Biodata handbook: Theory, research and use of biographical information in selection and performance prediction. (pp. 391-419). CPP Books: Palo Alto, CA.

[141] *See* 5 C.F.R. § 731.202(b) which explains that one of the grounds upon which an individual may be found unsuitable for employment includes making a "[m]aterial, intentional false statement, or deception or fraud in examination or appointment"); §§ 731.203-731.205 authorizes cancellation of eligibility; removal; cancellation of reinstatement eligibility; and debarment if an individual is found unsuitable.

[142] IAG-CTEA, (2009).

Warnings should be combined with an application attestation statement where applicants indicate, by signing or entering initials, that they have read the warning, understand it, and that everything they have said is true to the best of their knowledge. Warnings carry additional weight if applicants are asked to provide names and contact information for people who can verify the information they provide. Figure 3 provides a sample warning statement.

Figure 3. Sample Verification Warning

Warning:
All of the information you provide may be verified by a review of the work experience and/or education as shown in your application forms, by checking references, and through other means, such as the interview process. This verification could occur at any stage of the application process. Any exaggeration of your experience, false statements, or attempts to conceal information may be grounds for rating you ineligible, not hiring you, or for firing you after you begin work.
Applicant Attestation:
By checking the box to the left of this statement and by typing my full name in the space below, I declare and affirm that I have read and fully understand that:
1. Any misrepresentation or material omission of facts on this assessment questionnaire or in any other materials I submit in support of my candidacy (including but not limited to the application), or in any oral statements I may make during the selection process shall be sufficient cause to end further consideration of my candidacy.
2. Persons listed as having knowledge of my past accomplishments on this assessment questionnaire may be contacted for verification purposes.
3. An offer of employment is contingent on successful completion of the entire employment selection process, including the receipt and review of references, satisfactory to the agency.
4. This verification may, but need not, begin prior to my receiving an offer.
5. I will be evaluated only on the information submitted.

Verification by documentation. Another applicant-driven verification strategy is to request that applicants provide documentation of training and experience as part of their application materials. Requested documentation may include, a resume, certificates or professional license information, or documents from previous work experience. Such documentation may be used to verify applicant claims on an occupational questionnaire or other T&E assessment.

Some judgment is necessary to decide how much evidence is provided by each document. Documents from trusted sources, such as licenses or performance appraisals, should carry greater weight than those created by the applicant. However, even applicant-generated documents can be useful. For example, the contents of an applicant's resume could be compared with his or her claims on an occupational questionnaire for consistency. Any inconsistency discovered could reflect a mistake by the applicant or some form of misrepresentation. Either case could prompt a follow-up discussion with the applicant to reconcile the inconsistency, especially if the applicant appears otherwise highly-qualified for the position.

Verification by elaboration. A third applicant-driven verification strategy requires applicants to produce a written justification for their claim to have had experience with a task or to possess an ability.[143] Research shows less exaggerated responding when this type of verification is used. There is less inflation for the specific questions that are verified *and* for the other questions that appeared with them on the questionnaire.[144] Verification by elaboration works best when applicants do not know beforehand which questions will require elaboration. This encourages them to respond to all questions at a level which they can justify. Although this strategy can be used for several T&E assessments, it is most often used for questionnaire-based T&E assessments.

For example, after an applicant has completed self-ratings of their experience on an occupational questionnaire, a subset of the questions is selected for elaboration. Applicants are provided with a text box[145] to enter a brief justification of each task or competency rating. They are asked for additional details such as how often have they done a task, how much time it typically took, how much help they had, and what tools they used.[146] These narratives are later reviewed by experts in the jobs and tasks. Adjustments may be made to an applicant's ratings to better match the level of experience described and justified in the applicant's narratives. Applicants are told in advance that such adjustments may take place.[147] An example of such an elaboration narrative appears in Figure 4.

---

[143] Current OPM Hiring Reform regulations preclude use of these and other assessments requiring written narratives as part of the initial screen in Federal hiring.

[144] Schmitt, N. & Kunce, C. (2002). The effects of required elaboration of answers to biodata questions. *Personnel Psychology*, 55(3) 569–587; and Schmitt, N., Oswald, F. Kim, B. & Yoo, T. (2003). Impact of elaboration on socially desirable responding and the validity of biodata measures. *Journal of Applied Psychology*, 88(6), 979-988.

[145] This is easiest to implement in an online assessment system because the order in which applicants see information can be controlled. But it can also be done in a pencil-and-paper environment.

[146] IAG-CTEA, (2009). Also, Schmitt, N. & Kunce, C. (2002). The effects of requiring elaboration of answers to biodata questions. *Personnel Psychology*, 55(3), 569-587.

[147] See, for example, Boyce, A., Carter, L., Cober, A., Montanari, M. & Quinones, R. (2010). The practical implications of a narrative review process for mitigating the impact of socially desirable responding on self-report training and experience assessments. SIOP.

Figure 4. Verification by Elaboration

Sample Narrative (Specific Item):
Using the text box below, please provide a sample accomplishment from your education and/or experience that supports your response to the item above. Your sample accomplishment should be brief (a few sentences), specific, and verifiable.
Sample Narrative (Content Area, Certain Responses):
Select the choice that best describes your experience in working in or leading work-related teams.
A. I have had no experience in working in or leading work-related teams.
B. I have participated as a team member, receiving task assignments and project goals from my supervisor or team leader.
C. I have served as a team leader, communicating task assignments and project goals to team members and monitoring their progress.
D. I have led several teams, ensuring that project goals and deadlines were met, and I have provided feedback to team members on their performance.
If you chose 'C' or 'D' in the previous question, please give the name and email address of someone who can verify your experience as a team leader.

Keeping applicants honest may not be the only benefit of asking applicants to elaborate on their claims—these elaborations may actually improve their ability to self-evaluate. The process of elaboration encourages applicants to thoroughly consider their degree of experience, knowledge, and ability while answering a question. The greater effort required ensures a certain level of motivation on the part of the applicants. It can make the assessment of experience seem fairer to the applicants. It also makes it harder for applicants to out-and-out fabricate because of the need to come up with many details and make sure they all fit together believably.

Verification with bogus tasks. One agency-driven verification strategy involves inserting descriptions of *bogus tasks*—tasks that do not actually exist—into a T&E assessment to identify candidates who indicate they have performed them. This technique is most effective if applicants are informed that these bogus items are scattered throughout the questionnaire. It is best to use several such bogus items and to only take action against applicants who demonstrate a strong pattern of claiming experience with them.[148] Simply warning that bogus tasks are present also reduces faking and exaggeration.

---

[148] Fluckinger, C.D., McDaniel, M. & Whetzel, D. (2009). Review of faking in personnel selection. In Mandal, M. (Ed.) (2009) Emerging Frameworks in Science and Technology Personnel Selection and Recruitment. McMillian: New Delhi.

The bogus task strategy works well to identify unqualified and dishonest applicants who rate themselves high on all tasks to obtain a good score. Studies have found that in some cases as many as half of applicants have claimed to perform a bogus task.[149] Yet, claiming to perform a bogus task does not always reflect dishonesty. Sometimes it happens from not knowing the proper name for a task or equipment, or genuinely believing (mistakenly) that the bogus task refers to something that they have actually done.

Despite the allure of the bogus task strategy, it is very challenging to implement well. First, it is not easy to write an effective bogus task question. The question must clearly signal its bogus nature to honest applicants while seeming like a real task to dishonest applicants. Second, this strategy is not appropriate for all occupations or jobs. It works best in subject areas with highly technical content, dense terminology, and rapid change. Finally, it is not appropriate for all applicant pools. Some applicants have a strong negative reaction when told that there are bogus tasks on an occupational questionnaire. This may damage an agency's recruitment efforts and reputation.

Verification by later assessment. Another agency-driven verification strategy is to use additional assessments to verify information provided by applicants on earlier assessments. MSPB has recommended that agencies use multiple hurdles (sometimes called sequential hurdles)[150] in the assessment process for positions. "Using assessment tools in succession can make the assessment process even more effective in managing the candidate pool and narrowing the field of qualified candidates."[151] Although the primary purpose for sequencing assessments is to narrow the applicant pool efficiently, it also provides the opportunity to use later assessments to verify the results of earlier assessments, such as confirming an applicant's assertions on an occupational questionnaire.

For example, an initial T&E assessment, such as an occupational questionnaire, is used to obtain an estimate of each applicant's proficiency. Lower scoring applicants are removed from consideration. Higher scoring applicants receive an additional assessment. In this later hurdle, an assessment is administered that directly measures present ability level in the same area that was assessed by the occupational questionnaire. Yet, unlike the occupational questionnaire, the second assessment is proctored and measures applicant ability in real time, greatly diminishing the risk of

---

[149] McGonigle, T. & Curnow, C. (2007). Measures of training and experience. In Applied Measurement: Industrial Psychology in Human Resources Management. Lawrence Erlbaum Associates: Mahwah, NJ; See also Burns, G. & Christiansen, N. (2011). Methods of measuring faking behavior. *Human Performance, 24(4), 358-372*; Doll, R. (1971). Item susceptibility to attempted faking as related to item characteristic and adopted fake set. *Journal of Psychology*, 77, 9-16.

[150] Gatewood, R., Felid, H. & Barrick, M. (2010). Human Resource Selection. Cengage Learning: Independence, KY.

[151] MSPB (2006). Reforming Federal Hiring: Beyond Faster and Cheaper. MSPB: Washington, DC.

applicant dishonesty. Applicants who have exaggerated their abilities to get through the first screen will be screened out by the second assessment.

Verification by reference check. A third agency-driven strategy verifies applicants' claims about their training and experience through consultation with people who are familiar with applicants' behavior in these past situations. Applicants are asked to supply names and usable contact information for people who can corroborate their claims.[152] In this type of verification, most often done for accomplishment records, these references are called verifiers. These verifiers are often supervisors, coworkers, clients, or others who were present when the applicant was doing the work that they claim responsibility for. As with other types of verification discussed in this section, simply telling applicants that reference checking will occur reduces falsification and exaggeration on T&E assessments. For additional information on how to conduct sound reference checks, see MSPB's report on reference checking.[153]

In summary, these applicant-driven and agency-driven verification techniques, used singly or in combination, can reduce the impact of applicant misrepresentation and dishonesty on T&E assessments, thereby helping to improve the validity of T&E assessments.

Improvement Strategy 3: Use Assessments that Focus on Present Proficiency

A straightforward, but often overlooked, solution to the challenges presented by T&E assessments is to use alternatives. One such alternative is to use present-oriented assessments. Instead of using the past as a proxy for an individual's current level of proficiency in a particular area, present-oriented assessments examine the abilities and proficiencies that individuals currently possess. They often take the form of tests or other structured activities where the applicant's performance is scored against a standard. Further, these present-oriented assessments can provide insights into what proficiencies individuals currently have that give them the potential to learn new skills and develop existing ones over time.

Present-oriented assessments can predict future job performance more accurately than most T&E assessments which rely on the past as a proxy for current proficiency. If carefully and properly developed, present-oriented assessments are usually more effective tools. Further, OPM has recommended that agencies consider using alternatives to T&E assessment when the cost of a hiring error is high due to the mission-critical nature of the position; when positions are nontechnical and do not require specific education and expertise; or when there is a history of management

---

[152] IAG-CTEA, (2009).

[153] MSPB (2005). Reference Checking in Federal Hiring: Making the Call. MSPB: Washington, DC.

dissatisfaction with previous hires or referrals for the position.[154]   MSPB has also recommended that Federal agencies avoid over-relying on T&E assessments for hiring and promotion decisions.[155]   Of course, there will be circumstances where T&E assessment is the best practical alternative. Time, funding, and other resource constraints often preclude developing or purchasing such preferable present-oriented tools.   Further, T&E assessments are often the best way to assess some technical and specialized skills. And sometimes T&E assessments are expected by executives and other applicants who prefer not to take a "test." However, before choosing a T&E assessment, it is wise to consider some alternatives like those described below.

Present-focused structured interviews.   As discussed in Chapter 2, structured interviews are easy to develop and highly valid assessments with a solid track record in Federal hiring.[156]   They are in the "Very Useful" category for predicting future job performance (see Table 1). Present-oriented structured interviews can be one alternative to past-oriented structured interviews and other T&E assessments in some situations. By presenting applicants with a hypothetical scenario or problem and asking them how they would respond, present-oriented questions reduce challenges of recall, lack of perspective, and dishonesty. The applicant's performance is in the present, where it can be observed and evaluated, rather than in the past. Additionally, since the interviews are scored in real-time by trained raters (not the applicant), structured interviews avoid challenges inherent in applicant self-assessment.

Simulations.   Simulations or current work samples are another present-focused assessment with a predictive validity of .54—slightly better than structured interviews. This puts simulations in the "Very Useful" category for predicting future job performance (see Table 1). MSPB has studied how agencies can best use such assessments.[157] Chapter 2 noted that *past* work samples are T&E assessments similar to accomplishment records. A current work sample avoids uncertainties about an applicant's role in an accomplishment by requiring a work product to be produced under controlled conditions. Although there are complex simulations, such as aircraft flight simulators, simple and inexpensive simulations can also be effective assessments. Often job-relevant abilities can be assessed with a writing sample or in-basket exercise.

---

[154] IAG-CTEA, (2009).

[155] MSPB (2010). Making the Right Connections: Targeting the Best Competencies for Training. MSPB: Washington, DC, p. 47.

[156] MSPB (2003). The Federal Selection Interview: Unrealized Potential. MSPB: Washington, DC.

[157] MSPB (2009). Job simulations: Trying out for a Federal job. MSPB: Washington, DC.

A simulation may also serve as a realistic job preview that encourages some applicants to opt out of the selection process after discovering that they do not enjoy certain tasks required by the job. This benefits both employer and applicant by reducing the mutual problem of a poor person-job fit. It may also encourage applicants to drop out who may be prone to exaggerate or falsify information on subsequent assessments.

These two present-oriented strategies and others can be used instead of— or along with—T&E assessments across a number of occupations and to evaluate a range of competencies. Other present-oriented assessments may work better with particular competencies or occupations. We reiterate that hiring managers must consider advantages and disadvantages, and exercise wise judgment, when selecting an assessment for their particular hiring needs.

Improvement Strategy 4: Evaluate Developments in T&E Assessment

The ability of T&E assessments to predict job performance has increased over time as new types of assessments have been devised, and existing assessments have been improved. This is not only progress, but a reminder that further innovations may emerge. Work environments, policies, and practices may also change in ways that affect T&E assessment.

Those who use T&E assessments should watch for changes that affect the validity of these assessments. Here are a few of the changes that seem possible in the near future:

- Social networking and access to online work history information may change the way we do reference checking and other types of verification;

- Transparency and widespread sharing of documents on the web makes more written work products more readily accessible. This may change the way an applicant's previous accomplishments are verified;

- Changes in patterns of employment—such as an increase in the typical number of employers over the course of an employee's career— may make it more difficult to interpret applicant work history information;

- Changes in collaborative work may affect how prior work products are evaluated. Video-conferencing, document sharing, and collaborative software make it easier to work closely with others, perhaps making the level of involvement of each participant more difficult to verify;

- New ways of using automated testing platforms may become common practice, such as using applicant response patterns to determine what questions require verification; and

- There likely will be new ways of assessing training and experience, and new ways of determining the quality of T&E assessments. Watch for well-designed validity studies which assess the relationship between applicant scores on T&E assessments and their subsequent job performance.

Some relevant changes may already be under way. One interesting proposal is the use of *Skills Transcripts* which contain information about an individual's skills and accomplishments learned on the job and which could follow employees through their careers.[158] The idea is similar to the competency scores in OPM's USA Hire battery which can be used to apply for different jobs at different agencies.[159] The "skills transcript" idea has its challenges, including requiring great cooperation between employers. However, this and similar trends may change the way T&E information should be evaluated.

Training and experience assessments can bring value to a hiring or promotion process if implemented appropriately. As the previous chapters have discussed, there are advantages and disadvantages to using T&E assessments, and each hiring manager's threshold for the right balance will be different. Whenever possible, agencies should augment the usefulness of T&E measures through adopting one or more of the improvement strategies that are currently available. Further, agencies should keep a lookout for future developments in T&E assessment. Such developments can play a critical role in shaping the effectiveness of an agency's overall assessment approach, workforce composition, and ultimately mission success.

---

[158] Masie, E. (2011). Skills transcripts for workers? Learning TRENDS, May 22, 2011.

[159] www.opm.gov/services-for-agencies/assessment-evaluation/online-assessment/ . A similar single-score project is under way to create a post-college certification exam that could avoid current problems caused by non-comparability of grading practices across educational institutions. See http://t.nbcnews.com/business/not-enough-graduate-college-now-theres-exit-exam-8C11006596.

Chapter 5 - Conclusions and Recommendations

MSPB examined T&E assessments as part of its mission to uphold the Merit System Principles and support effective Federal workforce management. Additionally, T&E assessments are widely used, well regarded, and hiring decisions made using T&E assessments can have a lasting impact on the Federal workforce. Further, given the austere economic climate currently enveloping Federal agencies—and the resultant budget and resource constraints and limited hiring opportunities—it is all the more critical that hiring managers make wise decisions on which assessments they use to bring in talent. Although T&E assessments can be useful in Federal hiring, they also have disadvantages that require attention and mitigation. Together these factors made T&E assessments a prime candidate for study.

Previous chapters discussed the value of training and experience assessments; reviewed the main types of training and experience assessments; highlighted some challenges common to all such assessments; and provided methods for increasing the accuracy and usefulness of T&E assessments. This chapter summarizes this study's findings about training and experience assessment and presents recommendations intended to increase the effectiveness of these measures and improve Federal hiring practices.

Conclusions

1. **T&E assessments are worth thoughtful consideration by hiring managers, HR professionals, and agency decision makers.** T&E assessments fall within the scope of Merit System Principles that address selection, training, and effective management of the Federal workforce. T&E assessments are generally regarded as fair measures of job-related abilities and they are widely-used by Federal agencies. Some T&E assessments are used early in multistage assessments where they may have great impact on later hiring decisions. T&E assessments may also give an unintended advantage to internal applicants who have performed a job superficially similar to the one for which they are applying.

2. **There are professional standards that inform the development and evaluation of T&E** assessments for selection and promotion. The most important standard is high predictive validity: "The extent to which the assessment method has been shown to accurately measure a job-related competency and/or predict successful performance on

the job."[160] T&E assessments have a range of validity, with some being clearly superior to others. However, the validity of some T&E assessments is likely to increase over time as improved techniques emerge from research and practice.

3. Some T&E assessments are better predictors of future job performance than others. The better predictors of job performance include: direct questions, occupational questionnaires, accomplishment records, KSA narratives, and reference checks. Structured interviews and biodata, which can include T&E information, are also useful in gauging an applicant's likelihood of success in a job.

4. Some commonly-used T&E assessments do not predict future job performance well. They include years of job experience, number of training classes attended, grade point average, and resumes (when used without verification).

    T&E assessments focus on the past. This seems like a strength, but has disadvantages. It is difficult to determine what happened in the past; people differ in their abilities to recall and accurately report the past; and the requirements of the workplace change over time, sometimes making past experiences irrelevant to current duties.

5. Applicants are not good at evaluating their own abilities. They lack perspective on the range of their abilities and how their skills compare to others; they have difficulty recalling their past experiences accurately; and they often have not had the training to appropriately rate their proficiencies.

6. Some applicants exaggerate or deliberately misrepresent their training and experience. T&E assessments rely on applicants to supply information about their qualifications for a particular job. Unfortunately, some applicants are careless or dishonest in how they represent their past experiences.

7. T&E assessments can be improved. The recommendations section below contains several possible strategies to improve assessment of T&E.

---

[160] OPM's Assessment Development Tool at apps.opm.gov/ADT/content.aspxapps.opm.gov/ADT/content.aspx.

Recommendations

The conclusions above support the following recommendations:

1. Low-validity T&E assessments should not be used in hiring decisions. They include years of job experience, number of training classes attended, grade point average, and resumes (when used without verification). Any assessment which is not supported by validity evidence should also be avoided.

2. Strengthen T&E assessments with techniques that improve their accuracy. Successful techniques include grounding the assessment in a careful and recent job analysis, writing clear questions that aid recall of job-relevant events, crafting rating scales that ease the evaluation task, providing training in self-rating, and using scoring procedures that are professionally-developed.

3. OPM should develop brief self-rating training for applicants. Such training would improve accuracy of occupational questionnaire ratings by applicants for Federal jobs. Centrally funded, it would be available to applicants to all agency jobs, improving the hiring process Governmentwide.

4. Verify information supplied by applicants. Applicant information on T&E assessments can be corroborated through agency-driven and applicant-driven verification strategies. Agency-driven techniques check the veracity of applicant supplied information, while applicant-driven techniques encourage applicants to provide honest and accurate information.

5. Consider present-oriented alternatives to T&E assessment. Before deciding to use a T&E assessment, a hiring manager should consider using present-oriented assessments with higher predictive validity, such as present-oriented structured interviews or work samples.

6. Monitor developments. Researchers and practitioners will continue to improve T&E assessment. Work environments, policies, and practices may also change in ways that affect T&E assessment.

MSPB recommends that agency decision makers, hiring managers, and HR specialists consult this report before using a T&E assessment. This information, in combination with their best professional judgment, can help improve the assessment procedures in their agencies and the hiring or promotion decisions based upon them.

APPENDIX A - Understanding Validity

What is validity?[161]

When people talk about assessments, such as an employment examination, they often ask whether the test is valid or not. Validity is the extent to which a test (such as a psychological test, an academic aptitude test, or an employment selection instrument) measures what it claims to measure. In the context of hiring, validity is an estimate of the degree to which the result of the assessment (such as an applicant's score on an occupational questionnaire, or the rating received by an applicant's accomplishment) predicts future performance on the job.

Why does validity matter?

Validity matters for both practical and principled reasons. First, a test must be valid for its results to be accurately interpreted and applied. In hiring, an assessment must have acceptable validity for it to help an agency screen job applicants, or help a manager select among referred candidates. An assessment that has no validity (such as handwriting analysis) or low validity (such as years of experience) does not help a manager determine whether a candidate will perform successfully if hired, or identify the candidates who are best-qualified and most likely to perform at a high level.

Second, the merit system principles, which envision "selection solely on the basis of relative ability" implicitly require Federal agencies to use valid assessments to evaluate applicants and make hiring decisions. Regulations such as 5 CFR 300 and guidelines make this expectation explicit and prescribe measures to promote the use of job-related criteria and valid assessments.

How is validity measured?

The validity of a particular assessment method (such as the T&E measures discussed in this report) is measured through research that collects and analyzes data on (1) the scores that applicants with different levels of ability receive on the assessment and (2) their level of performance on the job.

The validity statistics presented in this report are based on research that synthesized results from a large number of research studies to determine the expected validity of each type of assessment—so long as that assessment is developed according to professional standards and current best practices. This kind of information can be used along with cost and other constraints to design an assessment approach for a particular type of job.

---

[161] Some of the material in this Appendix was adapted from an online Q&A on validity available at psychology.about.com/od/researchmethods/f/validity.htm.

What is a validity coefficient?

Validity is not determined by a single statistic, but by a body of research that demonstrates the relationship between an assessment and the behavior it is intended to measure. But a single statistic, the *validity coefficient,* is often used to summarize the validity evidence for a given assessment. Validity coefficients range between 0 and 1. Values closer to 1 indicate that an assessment is a good predictor of job performance; values closer to 0 indicate a poor predictor.

We cannot expect validity coefficients to approach 1.0 for many reasons which have nothing to do with assessment quality.[162] The results of past validity studies suggest how to set our expectations. Assessments with validity coefficients of .35 or above are considered very useful in predicting future job performance. A coefficient between .21 and .35 indicates an assessment that is likely to be useful. Assessments with validity coefficients between .11 and .20 may be useful in some circumstances. An assessment with a coefficient of .11 or below is unlikely to be useful.[163]

Can an assessment method's validity coefficient change?

The ability of an assessment to predict job performance can improve over time. This occurs as research and practice identify ways to improve the way they are developed, administered, and scored. The high validity of structured employment interviews that MSPB reported in a previous study[164] were achieved because such improvements now distinguish them from unstructured interviews.[165] MSPB's study of reference checking[166] noted that such a distinction between structured and unstructured reference checks was not yet made in validity studies. Subsequent validity studies found an improvement in reference checking which adopted best practices, relative to less effective unstructured reference checking.[167] T&E assessments can also benefit from this type of progress.

---

[162] Schmidt, F., Caplan J., Bemis, S., Decuir, R., Dunn, L. & Antone, L. (1979). The behavioral consistency method of unassembled examining (TM-79-21). Washington, DC: U.S. Office of Personnel Management, Personnel Research and Development Center.

[163] Sproule, C. (2009). Rationale and research evidence supporting the use of content validation in personnel assessment. IPAC. Available at www.ipacweb.org.

[164] MSPB (2003). The Federal Selection Interview: Unrealized Potential. MSPB: Washington, DC.

[165] Huffcutt, A. & Culbertson, S. (2011). Interviews. In S. Zedeck (Ed.), APA Handbook of Industrial and Organizational Psychology. APA: Washington, DC (Vol. 2, pp. 185-204); See also Huffcutt, A. & Arthur, W. (1994). Hunter & Hunter (1984) revisited: Interview validity for entry-level jobs. *Journal of Applied Psychology*, 79, 184-190.

[166] MSPB (2005). Reference Checking in Federal Hiring: Making the Call. MSPB: Washington, DC.

[167] Taylor, P., Pajo, K., Cheung, G. & Stringfield, P. (2004). Dimensionality and validity of a structured telephone reference check procedure. *Personnel Psychology*, 57(3), 745–772.

APPENDIX B - Validity Coefficients

The table below contains the predictive validity coefficient for each type of assessment discussed in this report.

How well do different types of assessment predict future job performance?		
Practical Value	**Assessment[167]**	**Validity**
Very Useful	*Simulation*	.54
	Structured Interview	.51
	Accomplishment Record	**.45**
Likely Useful	*Biodata*	.35
	Occupational Questionnaire	**15 to .28**
	Reference Checks	**.26**
	KSA Narratives	—
Possibly Useful	Years of Experience	.18
	Transcript/GPA	**.17**
	Resume (by itself)	**.11**
Not Useful	Number of Training Classes	—
Variable	Direct Question	—

[167] Training and experience assessments listed in this column are in **bold**. Assessment names in *italic* indicate assessments which contain T&E questions along with other types of questions.

Evaluating Job Applicants:

The Role of Training and Experience in Hiring