
Psychological Testing in Personnel Selection: Contemporary Issues in Cognitive Ability and Personality Testing

**Steven L. Thomas and **Wesley A. Scroggins*
Missouri State University

This paper examines the development of personnel selection testing from the late 19th century to the present, emphasizing general cognitive ability and personality testing. The development of methods and standards in employment testing is examined with particular emphasis on selection validity and utility. The issues of fairness and discrimination in cognitive ability selection testing are explored. The transformation of older models of personality into the current Big-Five personality paradigm is discussed. The utility and fairness of personality testing for modern organizations is explored, particularly when used as part of a composite selection process with cognitive ability testing.

Key Words: Personality testing, Cognitive ability testing, Selection testing, Test validation and utility

Introduction

It is widely recognized that many human resource functions have the capacity to dramatically alter the effectiveness of organizations. None have more potential impact on an organization's effectiveness and its ability to develop a sustainable competitive advantage than the staffing function. The role of human resources in creating competitive advantage has been broadly acknowledged. The resource-based view offered by Barney and Wright (1998) argue that human resource skills add value because talent is rare, nonsubstitutable, and difficult to imitate. Similarly, well known best practice models (Pfeffer, 1995) argue that traditional sources of competitive advantage such as economies of scale, proprietary technology, or protected markets have become less important in sustaining long-term competitive advantage than the manner in which companies utilize their human resources.

These views of human resources as a source of competitive advantage all contain a common thread. To achieve competitive advantage through people, organizations must be able to select individuals who have exceptional skills and whose talents, values, and motives best fit the organization's culture, structure, and reward systems. If it is true that talent is rare and vital to organizational success, the organization's system of selection must include processes that allow companies to accurately identify aptitude, ability, and other characteristics in applicants that are recognized as contributing to organizational effectiveness. This need underscores the pivotal role of the staffing function and the importance of psychological testing in the development of sustainable competitive advantage since it is, to a great extent, these instruments that allow an organization to identify desirable candidates.

If we can assume that this contemporary view of competitive advantage through people is a paradigm widely embraced by both managers and scholars, it follows that psychological testing of job applicants is likely to become more important in the future. It is imperative that managers understand the potential and the limitations of psychological testing in employee selection. To that end, this paper reviews many of the issues associated with the development and use of psychological testing in employee selection, specifically focusing on developments in two widely used sets of instruments: tests of cognitive ability and personality.

The Roots of Psychological Testing

The use of paper-and-pencil psychological tests in human resource selection was essentially nonexistent prior to the beginning of the 20th century. The contemporary application of psychological tests and measures to personnel selection can be traced to the dual influences of the turn-of-the-century industrial psychologists and the field of management science. Although the investigation of personality has roots that extend to the ancient Greeks, many psychologists of the late 19th century viewed the application of psychological testing to problems in business and industry with disdain (Hearnshaw, 1987). By the end of the 19th century, however, the field of industrial psychology emerged with individuals such as Walter Dill Scott and Hugo Munsterberg advocating the exploration of psychological principles to applied problems in education and business (Mankin, Ames, & Grodski, 1980). The field of industrial psychology and the role of psychological testing achieved a substantial level of legitimacy when in 1916 the National Academy of Sciences created the National Research Council, a group of prominent psychologists who developed a set of tests and measures to select and place troops during World War I (Driskell & Olmstead, 1989). Despite some reluctance within the military, the government funded the testing process and some 3.5 million soldiers were tested and placed, thus validating the role of psychological testing in organizations (Van De Water, 1997).

At about the same time, the influence of the industrial engineers provided additional impetus for psychological testing in selection. The influence of Frederic Taylor began with his late 19th century

writings addressing the problems of industrial efficiency by relying on the scientific analysis of work through time and motion studies, and on the scientific selection of workers that matched job characteristics and rewards to individual worker skills and abilities (Taylor, 1916, in Mankin et al., 1980). Taylor's successors, most notably Frank and Lillian Gilbreth, worked to refine the approaches of scientific management, especially in attempts to consider the psychology of the worker, and formed closer alliances with industrial psychology.

After World War I, American business grew in size and complexity and faced increasing competition and employment regulation. The natural response was a push for the development of rational management systems and the increasing application of scientific methods to organizational problems. A group of individuals referred to as the "entrepreneurial psychologists" expanded the field of industrial psychology through their marketing efforts and the establishment of professional organizations and journals (Van De Water, 1997, p. 487). Ultimately, the control of the field fell to the academic community who challenged many of the conventional tools of selection such as employment interviewing and character analysis, and began to develop psychological instruments to take their place. The application of the scientific method to selection saw standards for test development, evaluation, and validation emerge. The distinction between scientific management and industrial psychology became more pronounced as psychologists began to emphasize the importance of individual factors such as personality and intelligence rather than contextual factors such as incentives (Van De Water, 1997; Viteles, 1932).

The field of psychological testing continued to expand throughout World War II as the federal government established organizations such as the Committee on Service Personnel and Selection to investigate the role of psychological testing in the war effort. Throughout the war, psychologists continued advancing the application of psychological testing to selection, training, and performance evaluation (Driskell & Olmstead, 1989). The effectiveness of psychological testing during the war effort has been documented (Flanagan, 1947). As a result of these successes, several organizations were established to support research: the Office of Naval Research, the National Science Foundation, the Army

Research Institute for the Behavioral Sciences, and the Air Force Human Resources Laboratory. Psychologists continued the development of selection and classification testing culminating in the use of the Armed Forces Qualification Test and the Armed Services Vocational Aptitude Battery as widely accepted instruments for selection, placement, and training decisions for recruits (Driskell & Olmstead, 1989; Lubinski, 1996).

Early Issues in Psychological Testing

The fields of industrial psychology, engineering, and management merged to deliver the practical application of psychological testing to organizational problems, but not all forms of psychological testing enjoyed the same level of acceptance. While cognitive ability testing became broadly established and gained rather wide public acceptance, other types of testing, most notably personality testing, did not gain the same level of support. The validity of cognitive ability tests for predicting job skill acquisition and performance has been widely established, as has its economic value to an organization through the selection of superior job candidates (Schmidt & Hunter, 1998). However, the potential success of cognitive ability testing has been tempered by the universally recognized fact that these types of tests tend to discriminate against some minority groups (Sackett & Ellingson, 1997).

Personality tests, by contrast, have not traditionally enjoyed the same level of support and their use in employment selection is much more controversial. Many experts conclude that personality tests as used in personnel selection lack validity, are easily faked, and are generally unsuitable for pre-employment screening (Blinkorn & Johnson, 1990). Many of the problems in personality testing originate with historical controversies over how personality is defined, how personality traits are described and measured, and how traits relate to behavior. Prior to the development of the Big Five personality models, general agreement on these issues was lacking (Heneman, Judge, & Heneman, 2000). The *Handbook of Industrial and Organization Psychology*, in its 1976 chapter on personality, describes a confusing set of motivation models, trait theories, and personality instruments originating from Hippocrates and continuing to the 1960s. A list of more than 30 personality instruments includes brief

and long self-report measures, measures of values, vocational interest measures, and projective techniques, the range and breadth of which serve to underscore the problems in defining suitable personality measures for selection purposes (Hough, 1976). Many of these measures are clinical or developmental instruments inappropriately used in personnel selection. Others have not demonstrated sufficient reliability or validity to be adequate selection measures (Heneman et al., 2000). While studies show that there is fairly consistent agreement on sets of personality traits common in successful managers (Grimsley & Jarrett, 1975; Jackson, Peacock, & Holden, 1982), historical reviews of the research exploring the validity of personality testing have pessimistically concluded that personality testing has little utility (Guion & Gottier, 1966).

Contemporary Research in Psychological Testing

Despite these less than stellar reviews, recent research has far more room for optimism about the role of personality testing in selection (Heneman et al., 2000). The remainder of this article is devoted to exploring these trends and issues associated with both tests of cognitive ability and personality, and in discussing the role of each in contemporary human resource selection.

The Issue of Test Validity

The field of psychological testing has not been exempt from the influence of fads and the introduction of ineffective tools, particularly in the manner tests are used and test results interpreted. Professional psychologists have continuously urged caution in the employment testing arena (Dawes, 1994; Dunette, 1966; Dunette & Hough, 1990, 1991, 1992; Lubinski, 1996; Lykken, 1991). Concern over the application of scientific principles to human resource selection has proven to be well-founded as the field has struggled with both methods and outcomes in attempts to identify instruments that would satisfy the need for scientific rigor and the tests of acceptance and utility demanded by practitioners.

The validity of selection measures is fundamental to useful personnel selection practice (Cascio & Aguinis, 2005). The exact definition of validity varies depending on the types of selection

instruments used and the situation. The validity of cognitive ability and personality tests is defined as the degree to which scores can be used to infer one or more measures of individual performance. This process is called criterion-related validity and it involves collecting test score data from either job applicants (predictive validity) or current employees (concurrent validity) and calculating the correlation between those scores and some measure of job performance (the criterion measure). Greater validity is evidenced by a greater degree of correlation between the test scores (predictors) and the measure of job performance (criterion measures). It should be noted that any specific selection instrument can have different validities since performance can be defined in any number of ways (Cascio & Aguinis), including how long it takes an employee to learn a job, measures of job tenure, measures of work output or job performance, or employee attitudes. Each measure of performance might correlate differently with a specific selection test.

Establishing criterion-related validity has an additional purpose. Since selection testing will eliminate some job candidates, the organization must be able to demonstrate that an instrument is job-related, should it generate adverse impact by disqualifying a disproportionate number of protected group members. Because of this legal imperative, the methods for establishing validity evidence are regulated and described in the EEOC's *Uniform Guidelines on Employee Selection* (U.S. Department of Labor, 1999). Under the *Uniform Guidelines*, companies may conduct their own validity studies, but the process is time-consuming, costly, and depends on having large sample sizes in order to achieve reasonable results. Companies may also rely on evidence of validity generalization or that a commercially purchased test has transportability in its application. This may occur when the test is fair, and validity evidence suggests that it has proven to be valid for similar jobs requiring similar levels and types of skills and abilities.

The notion that selection tests have validities that generalize to other jobs and situations beyond those specifically tested for is one that, although widely accepted now, has not always been embraced. Prior to the 1970's, many industrial and organizational psychologists believed that selection instruments were situationally specific in that test validity varied

not only from job-to-job but also from location-to-location (Guion, 1965). The implication was that an organization would have to conduct a separate validity study for each specific situation to insure accuracy in testing. This would be difficult and costly, and impractical or impossible to accomplish. This prescription proved to be unnecessary because by the end of the 1970s, researchers found that virtually all of the differences in validity outcomes were produced not by actual differences in the validities of the tests, but by statistical and measurement error brought about because of small sample sizes (Schmidt & Hunter, 1998; Schmidt, Hunter, McKenzie, & Muldrow, 1979). Many earlier validity studies had been completed on sample sizes of fewer than 100 employees. In such small samples, much of the variation in both test scores and performance measures can be due to idiosyncratic fluctuations in the data (Ghiselli, 1966; Guion, 1965; Lubinski, 1996). By the late 1970s, analytic tools such as meta-analysis allowed researchers to statistically pool the data across studies, thereby eliminating much of the impact of sampling bias. Results of these studies supported the concept of validity generalization, eliminated much of the need to perform in-house validity studies, and provided evidence to support the application of commercially available selection tests validated on different populations (Schmidt & Hunter, 1998).

The Cognitive Ability Test in Human Resource Selection

Cognitive ability has been defined in various ways and there is still substantial disagreement among experts as to whether cognitive ability is a general ability (general intelligence) or a label for a set of more specific and distinct abilities. It is useful to think of cognitive ability as ability related to thinking, perception, reasoning, verbal, and mathematical skills. Measuring cognitive ability for selection purposes is among the easiest and least expensive of all selection tests. Commercial tests such as the widely used Psychological Corporation's Wonderlic Personnel Test are readily available, take only about fifteen minutes to complete, and cost less than \$5.00 per applicant. Based on meta-analysis results, cognitive ability tests appear to be among the most valid of all psychological tests and are valid for most occupations. While these tests are more valid for jobs of greater complexity and tend to do better at

predicting training criteria than long term job performance, cognitive ability tests generalize across organizations and jobs and have been shown to produce large economic gains for companies that use them (Gatewood & Feild, 1998; Heneman et al., 2000).

Cognitive Ability Testing and Fairness in Selection

Despite the apparent predictive validity and high utility offered by cognitive ability testing, few companies use them as selection tools. One reason for this is that cognitive ability testing has been demonstrated to produce group differences or adverse impact (Cleary, Humphreys, Kendrick, & Wesman, 1975; Hartigan & Wigdor, 1989; Wigdor & Gamer, 1982). In general, groups including Hispanics and African-Americans score lower than the general population while other groups including Asian-Americans score higher (Heneman et al, 2000; Lubenski, 1995). The visibility of legal challenges to cognitive ability testing began with the famous 1971 *Griggs v. Duke Power* case. In this case, the Supreme Court ruled that when a selection test produces adverse impact against protected group members the company must be able to defend it by showing that use of the test is a “business necessity” for the operation of the business. The courts have held narrow interpretations of business necessity that require companies to show that no other acceptable selection alternative exists (Sovereign, 1999). As a result, many companies abandoned cognitive ability testing.

The problem over group differences in psychological instruments has proven to be a vexing one for psychologists, and is particularly troublesome as it regards the demonstrable success of mental ability testing. The field of industrial psychology has struggled with the clash between ethics and cultural sensitivity and intellectual honesty in dealing with the issues of group differences (Kimble, 1994; Lubinski, 1996). The contributions of cognitive ability testing are mitigated by policies limiting the use of selection tools that produce differential outcomes across protected groups. Therefore, many experts argue that some validity must be sacrificed to reduce adverse impact. The *Uniform Guidelines* require that where two procedures are reliable and valid, the company should select the one that produces the lesser adverse

impact (Equal Employment Opportunity Commission, 1978). This puts the staffing professional in the difficult position of having to weigh validity against adverse impact. Often, validity is sacrificed because less valid selection procedures are selected to avoid the risk of discrimination charges (Gatewood & Feild, 1998; Hartigan & Wigdor, 1989; Maxwell & Arvey, 1993).

Employers using a valid selection test typically desire to use a selection strategy that is as efficient as possible since it has been shown that hiring employees as much as one standard deviation above the mean in ability translate into economic values of as much as 40 percent more than the average employee (Schmidt & Hunter, 1983; 1998). Usually, the most efficient means is to incorporate a top-down method of selection where the best scoring candidates are selected first. Where organizations are concerned with addressing affirmative action and balancing efficiency with social consciousness, there may be opportunity costs that impact the bottom line (Schmidt & Hunter, 1998).

The adverse impact inherent in cognitive ability testing has been addressed in several ways. One solution that emerged in the late 1970s was a practice adopted by the Department of Labor for employment testing called “race-norming” (Cascio & Aguinis, 2005). In race-norming, the differences in selection test scores across races is viewed as an empirical fact but the raw scores are converted to percentile scores within the racial group. In 1989, the National Academy of Sciences actually endorsed this practice, concluding that the moderate validities of the General Aptitude Test Battery produced selection errors that were more pronounced on minorities (Hartigan & Wigdor, 1989). Many experts felt that this process was unfair to non-protected individuals and this view was embraced by Congress when race-norming was banned in the language of the *Civil Rights Act of 1991*. The relative lack of opposition by the scientific community provided some evidence of the realization that although race-norming created larger minority applicant pools, it came at the expense of selection utility (Gottfredson, 1994).

A company wanting to comply with the *Civil Rights Act's* prohibitions against discrimination could attack the problem of adverse impact by adopting a multiple hurdle process to selection such as applying a pass-

fail technique to cognitive ability testing. A cut-score can be manipulated to insure an acceptable number of protected group members in the selection pool and a secondary selection tool can be used to fill job vacancies without showing discrimination. The problem of this approach is that it produces less than optimal selection outcomes. Moreover, the process does not absolve the employer of liability. Based on the 1982 decision in *Connecticut v. Teal*, a test that has a pass/fail score preventing a large portion of protected group members from going on to the next step in the selection process is a civil rights violation regardless of the ultimate hiring outcomes (Sovereign, 1999).

Another approach that aims to address group differences in selection test scores is banding. Testing experts acknowledge that since no test is perfectly reliable, small differences in test scores can be due to error and other artifacts and lack statistical significance. Banding is a way to address this issue by designating bands or narrow ranges of scores (Cascio & Aguinis, 2005). All scores within a band are assumed to be equivalent for decision-making purposes and the organization is free to select any candidate scoring within the band. The proponents of banding argue that it can reduce adverse impact at low cost to utility, but even the proponents of banding recognize that top-down approaches have better predictive ability (Cascio, Zedeck, Goldstein, & Outtz, 1995). Those opposed to banding place greater emphasis on the amount of validity that is sacrificed (Gottfredson, 1994). Others point out that using banding with highly reliable tests that do generate adverse impact (such as cognitive ability tests) produces substantial loss of utility with little actual reduction in adverse impact (Heneman et al., 2000).

Science, Business, and Government

The situation described above highlights the differences between psychology as pure science, the application of psychology to business problems, and the role of the government in protecting the rights of the individual. Psychology deals with the application of the scientific method in pursuit of truth. When the psychologist seeks application to business problems, the goals become more complicated, particularly when attempting to balance the organization's need for efficiency with social objectives dealing with

individual rights. The humanitarian concern for the application of the scientific method to business has a long historical base as evidenced by early opposition to scientific management (Peterson, 1990). Concern regarding psychological testing and the social and ethical problems it produces became of particular importance after 1964. This was due to the legal imperative created by the passage of the *Civil Rights Act* and subsequent court challenges. The ability to use powerful selection tools such as cognitive ability tests became more limited by the necessity to avoid adverse impact (Hartigan & Wigdor, 1989). On the other hand, those legal constraints placed on psychological testing have also been the subject of continued criticism. For example, Herrnstein and Murray (1996) argue that real differences in intelligence test scores exist across protected groups, that cognitive ability test results are highly valid, and that ignoring those differences is unscientific and economically irresponsible. Because of the need to balance science with fairness, the application of certain types of tests to employee selection, most notably cognitive ability tests, has been difficult.

The Role of Personality in Selection

While the utility of cognitive ability testing in selection has been broadly accepted, the utility of personality testing, until relatively recently, has not. Historically, research documenting the low predictive validity and the potential for invasion of privacy based on item content has made their application as selection instruments questionable (Hogan, Hogan, & Roberts, 1996). In addition, the lack of agreement regarding the components of personality, the many different types of personality instruments available, and the inappropriate application of clinical instruments to selection have contributed to the reluctance of many organizations to apply personality testing to employee selection.

While the bulk of the research before the 1990s was critical of personality testing, the continued search for alternative instruments to ameliorate the disparate impact produced by tests of cognitive ability has renewed interest in personality instruments (Schmidt, Ones, & Hunter, 1992). Some have suggested that pairing personality testing with cognitive ability testing may be one means to enhance validity while reducing adverse impact (Ryan, Ployhart, & Friedel, 1998). Advocates of cognitive ability testing are

aware that it is a highly valid predictor of job performance because individuals with high levels of cognitive ability appear to acquire job knowledge faster and better, leading to increased levels of performance. It is also likely, however, that some aspects of personality enhance individual ability to apply intellectual capacity while other personality traits limit its application (Kaufman & Lichtenberger, 1999). Some set of personality traits relating to the individual ability to be receptive, willing to receive and use information, and interact with others, may ultimately prove to be a moderating variable that allows one to fulfill the potential of his or her cognitive ability in a work situation.

The Development of a Useable Model of Personality: The Big-Five Model

Given that personality research is potentially one of the most useful approaches to enhancing selection validity and utility, why has it taken so long for personality testing to gain acceptance? Modern researchers point to the historical lack of an accepted definition of personality and little consensus regarding personality traits. Models of personality have ranged from Eysenck's two basic dimensions of personality to Cattell's 171 traits with an abundance of models in between (Dunnette, 1976). It has only been recently, with the development of sophisticated meta-analytic techniques, that researchers have been able to aggregate specific traits into broad definitions of personality that have allowed the prediction of broad behaviors that define job performance (Heneman et al., 2000). Since the early 1990s, estimates of the validity of personality testing have inched upward due to the development of factorial approaches to personality that have become known as the Big-Five personality dimensions that appear to be the core elements of personality assessment (Barrick & Mount, 1991). As the Big-Five model has become more accepted, interest in the use of personality measures in selection has increased.

Research delving into the components of personality testing goes back at least three-quarters of a century to the work of Thurstone in the 1930s, who may have been the first to identify five independent components of personality (Thurstone, 1934). Other researchers found different numbers of components. Cattell (1947), for example, described twelve core factors of personality. When subsequent research

examined Cattell's variables, only five factors were shown to be unique, and researchers throughout the 1980s and 1990s have generally confirmed the five factor structure (Digman, 1990; Digman & Inouye, 1986; Goldberg, 1992; John, 1990; McCrae & Costa, 1985, 1987; McCrae & John, 1992; Wiggins & Pincus, 1992).

Although agreement on the names and descriptions of factors is not complete, the Big-Five factors have been labeled as follows: Extraversion (Factor I); Agreeableness (Factor II); Conscientiousness (Factor III); Emotional Stability (vs. Neuroticism, Factor IV); Openness to Experience (Factor V) (Heneman et al., 2000). New instruments that assess the Big-Five include the Personal Characteristics Inventory (PCI; Mount & Barrick, 1995), the NEO Personality Inventory (NEO-PI; Costa & McCrae, 1985), and the Hogan Personality Inventory (HPI; Hogan, Hogan & Roberts, 1996). All three are self-report, paper-and-pencil measures that are relatively inexpensive and efficient for selection purposes.

These factors have been shown to have reliably predicted supervisors' ratings of job proficiency and training proficiency (Barrick & Mount, 1991; Tett, Jackson, & Rothstein, 1991). Comparisons of the validity coefficients of the intellect and agreeableness factors and the well accepted cognitive tests for selection purposes, indicate that these correlations approach each other (Hogan, Hogan, & Roberts, 1996). Integrity tests, U.S. army personnel selection instruments, and customer service measures contain facets of the Big-Five dimensions and have been found to have validity coefficients in the .33 to .50 range (McHenry, Hough, Toquam, Hanson, & Ashworth, 1990; Ones, Viswesvaran, & Schmidt, 1993). These validities approach the validity of cognitive ability tests and dispute the pre-1990s position that personality tests have little validity in personnel selection applications. Hogan, Hogan, and Roberts (1996) have stated that those who label personality tests in employment selection as having low validities and limited utilities are wrong.

Evidence for the utility of personality testing continues to increase as researchers identify the correlates of personality traits and the importance of these relationships for work organizations. The development of both better models and methods of analysis has facilitated the examination of the effects

of personality traits on attitudinal and behavioral variables of interest to organizations. In one of the earliest meta-analytic studies using the Big Five paradigm, Barrick and Mount (1991) found that conscientiousness was a significant predictor of job performance across each of the occupational groups included in the study. They also reported that extraversion was a significant predictor of success in managerial and sales positions. At approximately the same time, Tett, Jackson, and Rothstein reported moderate validities for the traits of agreeableness and openness to experience with job performance. Mount and Barrick (1998) examined the relationships between the Big Five personality traits, job proficiency, and training proficiency. They reported that conscientiousness was significantly related to both job proficiency and training proficiency. Extraversion was found to be significantly related to both job performance dimensions in both managerial and sales positions. They also reported that openness to experience and agreeableness were valid predictors of training proficiency across all occupations included in the study.

In a more recent meta-analysis, Judge and Ilies (2002) examined the relationships between the Big Five traits and performance motivation. Their results indicated that neuroticism was negatively correlated with performance motivation, especially for goal-setting motivation. They also found conscientiousness to be a significant predictor of performance motivation across three motivational perspectives (goal-setting, expectancy, self-efficacy). These meta-analytic studies provide evidence that personality traits are valid predictors of employee motivation and job performance.

Research also suggests that personality is related to career success. Judge, Higgins, Thoresen, and Barrick (1999) studied the relationships between the Big Five traits and job satisfaction, income, and occupational status, which they used as measures of career success. Similar to other studies, they found that conscientiousness was a valid predictor of all three measures of career success, while neuroticism negatively predicted income and occupational status.

Personality traits may have significant effects on the types of psychological contracts that employees form with the employer (Raja, Johns, & Ntalianis, 2004). Individuals high in neuroticism were more likely to

form transactional psychological contracts, but individuals high in conscientiousness were more likely to form relational contracts. Relational contracts were found to influence employee attitudes and behaviors, being related to higher levels of job satisfaction, affective organizational commitment, and fewer intentions to leave the organization. Individuals with high neuroticism and low conscientiousness were also more likely to perceive a breach of the psychological contract.

As the research cited above suggests, there is considerable evidence that personality is directly or indirectly related to individual cognitive, attitudinal, and behavioral variables that provide value to organizations. This evidence, coupled with better methodology and the availability of more construct valid measures of personality, has made the use of personality tests in personnel selection and development activities more common.

Personality Testing, Adverse Impact, and Incremental Validity

Researchers have advocated adding a personality test to an ability test as a means of enhancing validity while reducing adverse impact of the selection system. The assumption underlying this argument is that there are factors related to job performance other than cognitive ability and that using these factors to predict job success produces less adverse impact. If these alternative factors are included with cognitive ability in a selection battery, then adverse impact should be significantly reduced. Recent research studies have concluded that the addition of a predictor producing smaller group differences (i.e., personality test) to a predictor producing higher group differences (i.e., cognitive ability test) does not reduce the potential for adverse impact to the degree that is often expected (Bobko, Roth, & Potosky, 1999; Schmitt, Rogers, Chan, Sheppard, & Jennings, 1997). These studies reported that the addition of alternative predictors (personality test, interview, biodata) to cognitive ability measures in a selection battery reduced, but did not remove the potential for significant group differences and adverse impact. This reduction in adverse impact appears to only occur with the addition of two or three predictors. Beyond the addition of two or three predictors, there is little gain in the reduction of potential for adverse impact (Sackett & Ellingson, 1997).

Although the use of personality tests with measures of cognitive ability may not have the desired effects on reducing adverse impact, it appears that the addition of personality measures to measures of cognitive ability as a composite predictor results in significant incremental validity (Bobko et. al., 1999; Schmitt et. al.,1997). These studies found that the validity of predictor composites was highest when alternative predictors were used in combination with cognitive ability. Though this combination of predictors resulted in the highest predictive validity, the inclusion of cognitive ability with these alternative predictors increased the potential for adverse impact.

Conclusion

These findings explain the conflict for organizations that desire optimal prediction in selection processes but also want to avoid the negative effects that optimal prediction might have on protected groups. For optimal prediction, it is best to create a predictor composite that includes a measure of cognitive ability and an additional measure such as a personality test. This will enhance incremental validity and prediction to the degree that the composite predictors are uncorrelated and account for unique variance in the prediction of job performance. Under these conditions, the potential for adverse impact increases. For maximum reduction of adverse impact, a predictor composite should exclude cognitive ability and include other predictors with high correlations among them. This should result in minimum potential for adverse impact but will also result in decreased predictive and incremental validity due the increased common variance shared among predictors and common variance that the predictors share with the criterion variable.

Research supports the use of personality tests in addition to cognitive ability measures where both are valid predictors of job performance (Bobko et. al., 1999; Sackett & Ellingson, 1997; Schmitt et. al.,1997). The use of personality tests with cognitive ability tests can reduce the potential of adverse impact created by the use of the cognitive ability measure and increase the predictive validity of the selection process. Organizations must be aware that the inclusion of a personality test will probably not reduce adverse impact to the degree that they might expect. Potential for adverse impact in the selection

process will likely continue to exist. Organizations must make their own decisions regarding their use of these predictors. Decisions should be based on the value placed on validity maximization versus potential adverse impact creation in the context of organizational values, needs, and strategy. Since personality tests used alone can result in adverse impact, some argue it would be better to use them in combination with cognitive ability for maximum predictive validity. They argue that this would be more defensible in court due to the increased validity of the selection process (Bobko, et. al, 1999).

The rich history of the application of the scientific method in human resource selection has demonstrated that measures go through iterations shaped by the tools of the science, changes in the social, cultural, and political environments, and organizational need. The manner in which various instruments are received is subject to change based on changes in these forces. The history of cognitive ability and personality testing have witnessed those perceptual nuances. While cognitive ability testing fell out of favor because of social and regulatory pressures, personality testing has been refined and has emerged as a valuable management tool. Both types of testing will be the subject of continued refinement and will likely play a pivotal role in human resource selection for the foreseeable future.

**Steven L. Thomas is a professor of management at Missouri State University. He received his Ph.D. from the University of Kansas. He has published extensively in labor relations and human resource management. **Wesley A. Scroggins is an assistant professor of management at Missouri State University. He received his Ph.D. in management from New Mexico State University. He has published in the Journal of Applied Social Psychology, Public Personnel Management, and Employee Responsibilities and Rights Journal.*

References

- Barney, J. B., & Wright, P. M. (1998). On becoming a strategic partner. The role of human resources in gaining competitive advantage. *Human Resource Management, 37*, 31-46.
- Barrick, M. R., & Mount, M. K. (1991). The Big-Five personality dimensions in job performance: A meta-analysis. *Personnel Psychology, 44*, 1-26.

- Barrick, M. R., & Mount, M. K. (1993). Autonomy as a moderator of the relationship between the Big-Five personality dimensions and job performance. *Journal of Applied Psychology, 78*, 111-118.
- Blinkhorn, S., & Johnson, C. (1990). The insignificance of personality testing. *Nature, 348*, 671-672.
- Bobko, P., Roth, P. L., & Potosky, D. (1999). Derivation and implications of a meta-analytic matrix incorporating cognitive ability, alternative predictors, and job performance. *Personnel Psychology, 52*, 561-589.
- Cascio, W. F., & Aguinis, H. (2005). *Applied psychology in human resource management* (6th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- Cascio, W. F., Zedeck, S., Goldstein, I. L., & Outtz, J. (1995). Selective science or selective interpretation. *American Psychologist, 50* (10), 881-882.
- Cattell, R. B. (1947). Confirmation and clarification of primary personality factors. *Psychometrika, 12*, 197-220.
- Cleary, T. A., Humphreys, L. G., Kendrick, S. A., & Wesman, A. (1975). Educational uses of tests with disadvantaged students. *American Psychologist, 30*, 15-41.
- Connecticut v. Teal*, 102 S.Ct. 2525 (1982).
- Dawes, R. M. (1994). *House of Cards*. New York: Macmillan.
- Digman, J. M. (1990). Personality structure: emergence of the five-factor model. In M. R. Rosenzweig and L. W. Porter (Eds.), *Annual Review of Psychology, 41*, 417-440. Palo Alto, CA: Annual Reviews.
- Digman, J. M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology, 50*, 116-123.
- Driskell, J. E., & Olmstead, B. (1989). Psychology and the military research applications and trends. *American Psychologist, 44* (1), 43-54.
- Dunnette, M. D. (1966). Fads, fashions, and folderol in psychology. *American Psychologist, 21*, 343-352.
- Dunnette, M. D. (Ed.). (1976). *Handbook of Industrial and Organizational Psychology*. Chicago: Rand McNally.
- Dunnette, M. D. and Hough, L.M. (Eds.). (1990, 1991, 1992). *Handbook of Industrial and Organizational Psychology* (2nd ed., Vols. 1, 2, & 3). Palo Alto, CA: Consulting Psychologists Press.
- Flanagan, J. (1947). Scientific development of the use of human resources: Progress in the Army Air Forces. *Science, 105*, 57-60.
- Gatewood, R.D., & Feild, H.S. (1998). *Human Resource Selection* (4th ed.). Forth Worth: Dryden Press.
- Ghiselli, E. E. (1966). *The Validity of Occupational Aptitude Tests*. New York: Wiley.
- Goldberg, L. R. (1992). The development of markers of the big-five factor structure. *Psychological Assessment, 4*, 26-42.
- Gottfredson, L. S. (1994). The science and politics of race-norming. *American Psychologist, 49*, 955-963.
- Griggs v. Duke Power Co.*, 401 U.S. 424 (1971).
- Grimsley, G., & Jarrett, H. (1975). The relation of past managerial achievement to test measures obtained in the employment situation: methodology and results - ii. *Personnel Psychology, 28*, 215-231.
- Guion, R. M. (1965). *Personnel Testing*. New York: McGraw-Hill.
- Guion, R. M., & Gottier, R. F. (1966). Validity of personality measures in personnel selection. *Personnel Psychology, 18*, 135-164.
- Hartigan, J. A., & Wigdor, A. K. (1989). *Fairness in employment testing: Validity generalization, minority issues, and the general aptitude test battery*. Washington, DC: National Academy Press.
- Hearnshaw, L. S. (1987). *The Shaping of Modern Psychology*. London: Routledge & Kegan Paul.
- Heneman III, H. G., Judge, T. A., & Heneman, R. L. (2000). *Staffing Organizations* (3rd ed.). Burr Ridge, Illinois: Irwin McGraw-Hill.
- Herrnstein, R. J., & Murray, C. (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Simon & Schuster.
- Hogan, R., Hogan, J., & Roberts, B. W. (1996). Personality measurement and employment decisions questions and answers. *American Psychologist, 51*, (5), 469-477.
- Hough, H. (1976). Personality and personality assessment. In M. D. Dunnette, ed., *Handbook of Industrial and Organizational Psychology, 571-607*. Chicago: RandMcNally.
- Jackson, D. N., Peacock, A. C., & Holden, R. R. (1982). Professional interviewers' trait inferential structures for diverse occupational groups. *Organizational Behavior and Human Performance, 29*, 1-20.
- John, O. P. (1990). The "big five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of Personality: Theory and Research* (pp. 66-100). New York: Guilford Press.
- Judge, T. A., Higgins, C. A., Thoresen, C. J., & Barrick, M. R. (1999). The big five personality traits, general mental ability, and career success across the life span. *Personnel Psychology, 52*, 621-652.
- Judge, T. A., & Ilies, R. (2002). Relationship of personality to performance motivation: A meta-analytic review. *Journal of Applied Psychology, 87* (4), 797-807.
- Kaufman, A. S., & Lichtenberger, E. O. (1999). *Essentials of WAIS-III Assessment*. New York: John Wiley & Sons, Inc.

- Kimble, G. A. (1994). A new formula for behaviorism. *Psychological Review*, 101, 254-258.
- Lubinski, D. (1996). Applied individual differences research and its quantitative methods. *Psychology, Public Policy, and Law*, 2 (2), 187-203.
- Lykken, D. T. (1991). What's wrong with psychology anyway? In D. Chieccetti and W. Grove, (Eds.), *Thinking Clearly About Psychology* (pp. 3-39). Minneapolis: University of Minnesota Press.
- Mankin, D., Ames, R. E., Jr., & Grodski, M. A., Eds. (1980). *Classics of Industrial and Organizational Psychology*. Oak Park, Illinois: Moore Publishing Company.
- Maxwell, S. E., & Arvey, R. D. (1993). The search for predictors with high validity and low adverse impact: Compatible or incompatible goals. *Journal of Applied Psychology*, 78 (3), 433-437.
- McCrae, R. R., & Costa, P. T. (1985). Updating Norman's "adequate taxonomy": Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology*, 49, 710-721.
- McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 51, 81-90.
- McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60, 175-215.
- McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. A., & Ashworth, S. (1990). Project A validity results: The relationship between predictor and criterion domains. *Personnel Psychology*, 43, 335-354.
- Mount, M. K., & Barrick, M. R. (1995). *Manual for the Personal Characteristics Inventory*. Iowa City, IA: author.
- Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology Monograph*, 78, 679-703.
- Peterson, P. (1990). Fighting for a better Navy: An attempt at scientific management (1905-1912). *Journal of Management*, 16, 151-166.
- Pfeffer, J. (1995). *Competitive advantage through people*. Boston, MA: Harvard Business School Press.
- Raja, U., Johns, G., & Ntalianis, F. (2004). The impact of personality on psychological contracts. *Academy of Management Journal*, 47 (3), 350-367.
- Ryan, A. M., Ployhart, R. E., & Friedel, L. A. (1998). Using personality testing to reduce adverse impact: A cautionary note. *Journal of Applied Psychology*, 83 (2), 298-307.
- Sackett, P. R., & Ellingson, J. E. (1997). The effects of forming multi-predictor composites on group differences and adverse impact. *Personnel Psychology*, 50, 707-721.
- Schmidt, F. L., & Hunter, J. E. (1983). Individual differences in productivity: An empirical test of estimates derived from studies of selection procedure utility. *Journal of Applied Psychology*, 68, 407-415.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.
- Schmidt, F. L., Hunter, J. E., McKenzie, R. G., & Muldrow, T. W. (1979). The impact of valid selection procedures on work-force productivity. *Journal of Applied Psychology*, 64, 609-626.
- Schmidt, F. L., Ones, D. S., & Hunter, J. E. (1992). Personnel selection. *Annual Review of Psychology*, 43, 627-670.
- Schmitt, N., Rogers, W., Chan, D., Sheppard, L., & Jennings D. (1997). Adverse impact and predictor efficiency of various predictor combinations. *Journal of Applied Psychology*, 82, (5), 719-730.
- Sovereign, K. L. (1999). *Personnel Law*. Upper Saddle River, NJ: Prentice-Hall.
- Taylor, F. W. (1916). The principles of scientific management. Bulletin of the Taylor Society. In D. Mankin, R. E. Ames, Jr., and M. A. Grodski, (Eds.), (1980). *Classics of Industrial and Organizational Psychology*. Oak Park, Illinois: Moore Publishing Company, 15-28.
- Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, 44, 703-742.
- Thurstone, L. L. (1934). The vectors of mind. *Psychological Review*, 41, 1-32.
- US. Department of Labor Employment Training Administration. (1999). *Testing and Assessment: An Employer's Guide to Good Practices*. Washington, DC: Government Printing Office.
- Van De Water, T. L. (1997). Psychology's entrepreneurs and the marketing of industrial psychology. *Journal of Applied Psychology*, 82 (4), 486-499.
- Viteles, M. (1932). *Industrial Psychology*. New York: W. W. Norton.
- Wigdor, A. K., & Gamer, W. R. (1982). *Ability Testing: Uses, consequences, and controversies*. Washington, DC: National Academy Press.
- Wiggins, J. S., & Pincus, A. L. (1992). Personality: Structure and assessment. In M. R. Rosenzweig, & L. W. Porter (Eds.), *Annual Review of Psychology*, 43, 473-504. Palo Alto, CA: Annual Reviews.