

FOCAL ARTICLE

---

# Personality Testing and Industrial–Organizational Psychology: Reflections, Progress, and Prospects

---

LEAETTA M. HOUGH  
*The Dunnette Group, Ltd.*

FREDERICK L. OSWALD  
*Michigan State University*

## Abstract

As the title suggests, this article takes a broad perspective on personality as it is conceptualized and measured in organizational research, and in the spirit of this Society for Industrial and Organizational Psychology journal, we framed the article as a series of 7 questions. These 7 questions deal with (1) personality and multidimensional models of performance, (2) personality taxonomies and the five-factor model, (3) the effects of situations on personality–performance relationships, (4) the incremental validity of personality over cognitive ability, (5) the need to differentiate personality constructs from personality measures, (6) the concern with faking on personality tests, and (7) the use of personality tests in attempting to address adverse impact. We dovetail these questions with our perspectives and insights in the hope that this will stimulate further discussion with our readership.

We introduce this article on the use of personality variables in industrial–organizational (I–O) psychology and personnel selection in particular by stating—with utmost confidence—what we will *not* be doing. This article is not another review or meta-analysis of validity findings for personality tests, as there are many current informative ones (e.g., for summaries of many meta-analyses, see Barrick, Mount, & Judge, 2001; Bartram, 2005; Hogan & Holland, 2003; and Hough & Furnham, 2003). This

article makes no attempt to resolve the insidious problem that job applicants might distort their responses—and even lie—on personality tests. Although we do draw on the meta-analysis literature, and we do provide some thoughts and concerns about faking personality tests, our more general goal is to provide a current perspective on personality testing and its use in organizational research and practice. In the spirit of this journal’s goal for stimulating conversation among all interested parties, the format of our article is to offer a series of questions about personality testing, dovetailed with our thoughts, to spur constructive discussion and future research and practice in I–O psychology. We hope that you find that this approach is valuable: not too abstract, not too concrete, and not too naïve.

---

Correspondence concerning this article should be addressed to Leaetta M. Hough. E-mail: leaetta@msn.com  
Address: The Dunnette Group, Ltd., 370 Summit Avenue, St. Paul, MN 55142.

Leaetta M. Hough, The Dunnette Group, Ltd.;  
Frederick L. Oswald, Department of Psychology,  
Michigan State University.

## A Bit of History—How We Got Here

Since the early 1990s, there has been a near-literal explosion of interest involving personality constructs in organizational research and practice. Prior to this was a two-decade lull in personality research that was an aftermath of the work in the 1960s by Guion and Gottier (1965), whose research concluded that personality variables have little or no systematic relationship to work-related criteria, and by Mischel's (1968) book that was interpreted as a full-scale attack on trait psychology. So what caused the resurgence of interest?

Project A, a project sponsored by the U.S. Army and conducted in the 1980s, was perhaps the most impactful personnel selection research project in our field in the second half of the 20th century. In addition to addressing the critical personnel selection issues at hand, it did nothing short of reorienting our field's thinking about both the predictor and the criterion space in a more sophisticated way that was multidimensional and construct oriented in nature. Regarding the measurement and modeling of personality variables, this meant incorporating the taxonomic thinking of the field of personality at that time. Thus, when the obligatory literature review on personality was undertaken for Project A, the research team contrasted the Guion and Gottier (1965) conclusion that operationally, personality variables were not useful predictors of work behavior or performance, with the very different conclusion of Ghiselli (1966) concerning the usefulness of personality variables in making personnel selection decisions. Ghiselli had indicated that he had included only personality scales that were conceptually appropriate for the job or work, but unfortunately, he did not describe his underlying rationale. Similar to Ghiselli, however, the Project A research team found that when they summarized the literature relating personality variables to job-related criteria according to *both* personality constructs and work-related criterion constructs taken together, then meaningful validity coefficients emerged (Hough, Eaton, Dunnette, Kamp, & McCloy, 1990). Note that the focal

constructs for Ghiselli and for Project A were not the five constructs in the five-factor model (FFM) of personality.<sup>1</sup>

In addition to this major literature review, Project A also included significant predictor and criterion development efforts as well as a large concurrent validation study. The conclusions from those activities were that when personality measures are based on theory-relevant constructs, and when these measures are administered to job incumbents and correlated with scores on job analysis-informed performance measures, then personality variables tended to predict discretionary "will do" criteria better than cognitive ability measures (McHenry, Hough, Toquam, Hanson, & Ashworth, 1990).<sup>2</sup> Importantly, these empirical relationships were consistent with the conclusions from the Project A literature review and Ghiselli's previous empirical findings.

Project A, together with earlier research by personality psychologists such as Harrison Gough, Robert Hogan, Douglas Jackson, Paul Costa, Lew Goldberg, Robert McCrae, and Auke Tellegen, set the stage for the Barrick and Mount (1991) meta-analysis of validities between Big Five personality variables and job performance, an article that became the most cited *Personnel Psychology* article in the 1990s (Hollenbeck, 1998).

We could spend time debating the advances and retreats in our field that have resulted since (see Hough, 2001, for a discussion of such advances, and see the exchange between Morgeson et al., 2007a, 2007b; Ones, Dilchert, Viswesvaran, & Judge, 2007; and Tett & Christiansen, 2007, for a stand-off on personality testing in selection settings). Instead, we prefer to focus more broadly on how our knowledge to date might improve our future understanding of individual behavior in work settings and to

- 
1. These constructs did not duplicate the FFM constructs. For a more complete description of the approach and results, see Kamp and Hough (1986). See Schneider and Hough (1995) and Hough and Schneider (1996), for a history of the FFM and its origins.
  2. See Campbell and Knapp (2001), for a thorough description of the activities and findings of Project A.

stimulate an exchange of ideas that will aid researchers and practitioners who apply our science in the world of work.

Although personality traits are challenging to conceptualize and measure, arguably more so than for **cognitive ability where the g factor drives much of the prediction**, we want to present a sample of some of the major findings that should encourage personality researchers in I–O psychology to press on.

### *Personality and Major Life Outcomes*

- Personality variables predict mortality, divorce, and occupational attainment. Correlations tend to be higher than those for cognitive ability predictors (meta-analysis of only prospective longitudinal studies: Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007).
- **Conscientiousness and its facets predict health behaviors, drug use, and mortality** (Roberts, Chernyshenko, Stark, & Goldberg, 2005; meta-analysis: Bogg & Roberts, 2004).
- Personality measures predict alcoholism (meta-analysis: Cooper-Hakim & Viswesvaran, 2002).

### *Personality and Performance*

- Personality variables predict overall job performance, objective performance, and getting ahead (meta-analyses: Barrick et al., 2001; Dudley, Orvis, Lebiecki, & Cortina, 2006; Hogan & Holland, 2003).
- Personality variables predict task performance (meta-analyses: Dudley et al., 2006; Hurtz & Donovan, 2000).
- Personality variables predict training performance and learning and skill acquisition (meta-analyses: Barrick & Mount, 1991; Barrick et al., 2001; Colquitt, LePine, & Noe, 2000).
- Personality variables predict contextual performance, such as organizational citizenship, altruism, job dedication, interpersonal facilitation, and generalized compliance (meta-analyses:

Borman, Penner, Allen, & Motowidlo, 2001; Dudley et al., 2006; Hurtz & Donovan, 2000; LePine, Erez, & Johnson, 2002; Organ & Ryan, 1995).

- **Personality variables predict overall managerial effectiveness, promotion, and managerial level** (meta-analysis: Hough, Ones, & Viswesvaran, 1998).
- Personality variables predict leader emergence and effectiveness as well as transformational leadership (meta-analyses: Bono & Judge, 2004; Judge, Bono, Ilies, & Gerhardt, 2002).
- Personality variables predict expatriate performance (meta-analysis: Mol, Born, Willemsen, & Van Der Molen, 2005).
- Personality variables predict goal setting and, conversely, procrastination (meta-analyses: Judge & Ilies, 2002; Steel, 2007).
- Personality variables predict creativity and innovation (Hough, 1992; meta-analyses: Feist, 1998; Hough & Dilchert, 2007).
- Personality-based integrity tests predict overall job performance (meta-analysis: Ones, Viswesvaran, & Schmidt, 1993).

### *Personality and Counterproductive Work Behaviors*

- Personality variables predict counterproductive work behavior (CWB) (meta-analysis: Berry, Ones, & Sackett, 2007).
- Personality-based integrity tests predict CWB (meta-analysis: Ones et al., 1993).
- Personality-based integrity tests predict absenteeism (meta-analysis: Ones, Viswesvaran, & Schmidt, 2003).

### *Personality and Team Performance*

- Personality variables predict team performance (meta-analysis: M. A. G. Peeters, Van Tuijl, Rutte, & Reymen, 2006).
- Personality variables predict getting along and teamwork (meta-analyses: Barrick et al., 2001; Hogan & Holland, 2003).

## Personality and Job Satisfaction

- Personality variables predict job and career satisfaction (meta-analyses: Judge, Heller, & Mount, 2002; Ng, Eby, Sorensen, & Feldman, 2005).
- Personality variables are highly correlated with subjective well-being with correlations up to .62 observed (.80 disattenuated) when substantively similar measures are grouped (meta-analysis: Steel, Schmidt, & Shulz, 2008).

## Mediated Models Involving Personality

- The relationship between conscientiousness and learning is mediated by goal commitment (Klein & Lee, 2006).
- The relationships between personality variables and work-related criteria are affected by motivational and self-regulatory mechanisms (Barrick, Mount, & Strauss, 1993; Erez & Judge, 2001; Kanfer & Heggestad, 1999; F. K. Lee, Sheldon, & Turban, 2003).
- Positive emotions moderate stress reactivity and mediate recovery from stress. Over time, positive emotions help people high in psychological resilience effectively recover from daily stress (Ong, Bergeman, Bisconti, & Wallace, 2006).
- Status striving mediates the relationship between extraversion and sales performance, such that high extraverts are high-status strivers, and high-status strivers are better performers (Barrick, Stewart, & Piotrowski, 2002).

If you view this research as a large expenditure of interest, energy, time—and yes, money—as well as a zero-sum game with attendant opportunity costs (i.e., focusing on some areas might mean focusing less on other areas), then we assert that it is critical for experts in the field to spend more time together discussing where such resources might be strategically allocated to address organizational issues via research and practice (see a similar view in J. P. Campbell, 1990). To that end, we present seven questions to generate discussion.

## 1. Given the development of multidimensional models of job performance, can refined measures of personality improve the understanding and prediction of work behavior?

The 1990s reflected rampant enthusiasm for research and practice involving personality measures used for personnel selection and other organizational practices, and at the same time, the 1990s also ushered in an era of creating multidimensional models of the job performance domain. **Multidimensional models of job performance have, by design and intent, included dimensions of performance that extend beyond core task performance.** A multidimensional model of general job performance across jobs includes **dimensions such as demonstrating effort, maintaining personal discipline, and facilitating peer and team performance** (J. P. Campbell, McCloy, Oppler, & Sager, 1993). Other multidimensional performance models or frameworks address **organizational citizenship behaviors (OCB) such as personal support, organizational support, and conscientious initiative** (e.g., Borman, 2004) and CWBs such as theft, unsafe behaviors, and absenteeism (e.g., Sackett & DeVore, 2002), and subsequent research has meta-analyzed or otherwise investigated the relationships between OCBs and CWBs (e.g., Dalal, 2005; Sackett, Berry, Wiemann, & Laczko, 2006). Note that because many if not most of these performance dimensions largely reflect discretionary behaviors (i.e., “will do” behaviors), they tend to relate more strongly to personality constructs than to those task performance dimensions that more heavily reflect competence or ability-based behaviors (i.e., “can do” behaviors).

Again, to reiterate, there has been a useful convergence in organizational research trends, where the resurgence of personality as a predictor of performance coincided with the theoretical expansion of the job performance domain into types of behaviors that are more discretionary or volitional and therefore more personality relevant. Adding to this coincidence is the finding that **personality variables not only show incremental validity but also can have some effect**

in reducing adverse impact (Hough, Oswald, & Ployhart, 2001), as we will discuss.

Given that these important refinements in the performance domain have been relatively recent, it perhaps should not be a complete surprise to us that it has taken time for measurement and modeling in personality testing to catch up by considering personality predictors at commensurate levels of refinement as the criterion. The idea of commensurate measurement is not new at all; it is in line with the tenets of **Brunswik's Lens Model** (Brunswik, 1952), which indicates that criterion-related validity will be higher when trait-relevant cues are more consistent with performance-relevant cues (setting aside other empirical concerns such as range restriction and differential measurement unreliability).

What is the most theoretically informative and practically useful level of commensurate measurement? Some have essentially argued that when predicting job performance, extremely broad measures of personality (e.g., integrity, core self-evaluation) yield the maximum criterion-related validity. Certainly, it is true that if the bandwidth of the criterion is wide, the most effective predictor will be a predictor that is matched to the criterion in terms of relevance and bandwidth. That is, if the **criterion is complex, the most effective predictor composite is also complex** (Hough & Ones, 2001). However, for breadth that is most informative, theory and data regarding relationships between narrow predictor and criterion constructs will help us more effectively understand validities for complex or aggregated criteria. Factor models and meta-analytic results of validities between homogeneous predictor constructs and homogeneous criterion constructs are informative to this end (Hough, 2001). To illustrate this claim, several studies (Hough, 1992; Hough et al., 1998; Roberts et al., 2005; Stewart, 1998; Tett, Steele, & Beauregard, 2003) taken together indicate that **facets of conscientiousness show meaningful differential prediction with outcomes relevant to organizations**. Hough et al. (1998) meta-analyzed validities of FFM personality factors and facets with overall man-

agerial job performance, finding that the observed meta-analytic validity of extraversion was .05, whereas the observed validities of its facets were .16,  $-.01$ , and .12 for dominance, sociability, and energy level, respectively. In a similar vein, Rothstein and Goffin (2006) reported 11 studies with higher validities for narrow compared to broad measures. Other research also supports higher or more interpretable predictive validity for narrower facets than for broader factors of personality (Ashton, 1998; Mershon & Gorsuch, 1988; Paunonen, 1998; Paunonen & Ashton, 2001; Schneider, Hough, & Dunnette, 1996).

A recent meta-analysis of conscientiousness–performance relationships (Dudley et al., 2006) also indicated that patterns of criterion-related validities at the facet level were theoretically informative. For instance, achievement predicted task performance for sales jobs (meta-analytic  $r = .26$ )<sup>3</sup> but order did not ( $r = .03$ ). Across jobs, dependability predicted counterproductive work performance ( $r = -.34$ ) but achievement did not ( $r = .00$ ). These results indicate that studies aggregating broad predictors and criteria potentially obscure our theoretical understanding and the prediction of practical outcomes (Hough & Schneider, 1996). Conversely, information on narrower predictor–criterion relationships has the potential to be more informative (assuming adequate sampling, measurement, statistical power, and so on), and practically speaking, it makes testing time more focused and effective by not including content from facets that are less theoretically relevant or predictive.

Related to the breadth of a construct, there has been much recent discussion in the organizational literature about *compound traits* in personality. **A compound trait can be defined as a set of trait constructs**; for instance, the compound trait of integrity

3. Validities were corrected for unreliability on both the predictor and the criterion, but the predictor reliabilities were generally high (estimated to be around .80), and thus, these corrections tend to reflect low criterion reliability (estimated to be around .60).



has been defined as a combination of conscientiousness, emotional stability, and agreeableness measures based on the Big Five (Ones, Viswesvaran, & Dilchert, 2005). The previous concern regarding commensurate measurement is at play here: The criterion-related validities for integrity may be meaningful, but the patterns of correlation between the component traits and any particular criterion may be obscured. For example, the compound trait of integrity might yield the same validity for two different criteria; yet, its validity might be driven by conscientiousness when predicting attendance and by emotional stability when predicting customer service.<sup>4</sup> Alternatively, a compound trait can reflect a subset of items sampled from a set of trait measures. Take, for instance, the measure of core self-evaluation (Judge, Erez, Bono, & Thoresen, 2003). It contains 12 items with content that overlaps the constructs of self-esteem, generalized self-efficacy, emotional stability, and locus of control. An issue with this sort of compound trait is that the items potentially sacrifice the construct validity of the factors from which they arise because there are very few items (if any) that strongly represent any one factor, and thus, the criterion-related validity of the measure might be driven by item-specific covariances as much as by variance because of the compound trait measure as a general factor.

Taken together, both of these types of compound traits have the potential to overlook the important details underlying them. Although compound traits may exhibit reasonable or even substantial prediction (as can any unit-weighted composite of measures or items), it can obscure interesting theoretical relationships that involve its constituent components, just as a general job performance measure can obscure important multidimensional information on the criterion side. Hough and Furnham (2003) summarized meta-analytic validities of

some finer grained constructs for predicting criterion constructs as well as for predicting overall work-related criteria. Their tables of meta-analyses organized by predictor construct and criterion construct show clearly that criterion-related validities vary widely depending upon the nature and breadth of both the predictor and the criterion construct, the type of job, and the relevance of the predictor for the criterion. Their tables also show the advantage of combining narrow constructs found to be relevant to a particular criterion construct into a compound variable that is found to correlate more highly with relevant criteria than any of the factors in the FFM. These sorts of compound variables are driven by careful examination of criterion-related validities; they are not items or traits combined solely on the basis of their own intercorrelations.

To be clear, we are not suggesting a level of refinement in personality measures that is impractical. What we are saying is that, it may ultimately be *more* practical to make meaningful refinements of personality variables so that relevant variables are more carefully aligned with theoretical models that include our major dependent variables of interest: task and contextual performance, training performance, team performance, job commitment, turnover, and job satisfaction (see Hough, 1992, 1998b; Hough & Schneider, 1996; Kanfer, Ackerman, Murtha, & Goff, 1995; and Tett & Christiansen, 2007).

We urge I–O psychologists to develop better databases that contain descriptive statistics, correlations, and other relevant effects concerning narrow personality constructs and criterion dimensions. Ideally, such databases would contain personality and criterion information separated by measurement method, situation, and other known moderators<sup>5</sup> of the relationship (see Hough, 2001). Data such as these are invaluable for meta-analysis undertakings and can help establish more systematic and

4. Remember that conscientiousness and emotional stability contain facets and also might be considered compound traits. But the example contains the spirit of the argument.

5. See Schneider and Hough (1995) and Hough and Schneider (1996), for a discussion of known moderators of the relationship between personality constructs and criterion constructs.

important relationships between personality and performance. Such a database was attempted in the Validity Information Exchange efforts of the 1950s and 1960s, but in the era of meta-analysis, the lack of such a database is virtually inexcusable for I-O psychology as a science.

## **2. How and why might personality research in I-O psychology be useful in extending its insights beyond the FFM?**

Useful taxonomies are important milestones in the history and advancement of a science. An adequate taxonomy is important for communicating efficiently, effectively revealing patterns of relationships, reaching a deep understanding of research results, and generalizing results meaningfully to other contexts (Fleishman & Quaintance, 1984; Hough & Ones, 2001). As we mentioned previously, taxonomic structure of variables—both for predictors and criteria—is important for discovering and understanding relationships hitherto obscured as well as for replicating those relationships in future research and applications.

Over 50 years ago, Cronbach and Meehl (1955) published their classic article on construct validity of psychological tests and their nomological nets. Following this, Guion (1961), Dunnette (1963), Smith (1976), and others argued for more sophisticated construct-oriented thinking about variables in our field. More recently, since Barrick and Mount's (1991) influential meta-analysis of the validities of FFM personality constructs for predicting work-related variables, most I-O psychologists have embraced the FFM as the taxonomic structure for personality.

Many researchers have concluded that the FFM of personality is robust and generalizes across different rating sources and cultures (e.g., Hogan & Ones, 1997; Saucier & Goldberg, 1998; Wiggins & Trapnell, 1997).<sup>6</sup> Although we harbor some criticisms of the FFM, we certainly do not want to lose the organizing benefits that broad-level taxonomies such as the FFM afford. We do not

want to return to the "good old daze" (Hough, 1997, p. 233) of hundreds of personality scales, each supposedly measuring unique variables.

We do acknowledge that the FFM has been important in understanding personality and provides an organizing structure for the myriad of personality scales. Nonetheless, we suggest that the breadth of the five factors of the FFM may be one reason why observed criterion-related validities are low and are criticized by some I-O psychologists as such (e.g., Morgeson et al., 2007a, 2007b). We do argue that even with low validities, **personality tests remain useful in selection because they provide incremental validity over cognitive ability, and even modest amounts of validity can translate into significant amounts of utility to the organization when aggregated across individuals and over time** (conversely, not administering a personality measure means losing this utility). But personality research shows that we can do better than the FFM in attempting to improve both our theories and our predictions.

Publications by Hough and her colleagues (Hough, 1992; Hough & Ones, 2001; Hough & Schneider, 1996; Schneider & Hough, 1995) have summarized research that has identified several additional constructs beyond the five in the FFM. More recently, K. Lee, Ashton, and deVries (2005) also added an **honesty-humility factor**, evidence that indicates that the FFM needs to be expanded.

The FFM of personality certainly has its critics. Several highly respected personality psychologists (e.g., Block, 1995; Loevinger, 1994; McAdams, 1992; Pervin, 1994; Tellegen, 1993) criticize the FFM as *theoretically* inadequate. For example, the theories and constructs of **moral development** (see Kohlberg, Levine, & Hewer, 1994) and ego development (Loevinger, 1966, 1976) can make important contributions to our understanding of individual behavior and performance in organizations. Similarly, constructs such as self-regulation and ego depletion (Baumeister, Gailliot, DeWall, & Oaten, 2006) can make important contributions to our understanding of individual behavior

6. However, see Hough and Schneider (1996), for a critique of this research.

and performance in organizations. Hough (1997) provided general criticisms of the FFM as follows:

- FFM overlooks important personality constructs; it is not comprehensive.
- FFM combines constructs that are better left separate.
- FFM is method bound and dependent upon the factor analysis of lexical terms.
- FFM consists of constructs so broad and heterogeneous that accuracy of prediction is sacrificed.

Typical efforts to identify the taxonomic structure of personality variables such as the FFM involve some type of factor analysis that only involves the scales themselves. Hough and Ones (2001) argue that a better approach for developing a useful taxonomy is to create profiles of relationships between a target personality variable and other psychological variables, such as job performance variables and other individual difference variables, such as those from the ability, motivation and interest domains. They labeled the approach “nomological-web clustering.” Profiles across target variables can be cluster analyzed by taking either a rational or mathematical approach, then “the pattern of relationships that a variable has with other variables should be similar to other variables that are in the same taxon or construct” (Hough & Ones, 2001, p. 237). This is a much more stringent approach to understanding whether a set of personality variables functions similarly and is in line with McDonald (1999), who said that measures claimed to be similar should exhibit similar correlational patterns with external variables.

Hough and Ones (2001, Appendix) used nomological-web clustering to develop conceptually similar personality constructs. We are unaware of any research that has followed up on this approach; yet, we believe work along these lines would be productive in helping us develop a better set of trait-based personality constructs for our field.

**3. Are situations important and, if so, how can we incorporate the situation or**

**context into measurement of personality variables?**

Hattrup and Jackson (1996) urge I–O psychologists to learn more about individual differences by understanding their function within the context of situations and their salient characteristics. Hough (2003) optimistically reported that Hattrup and Jackson’s plea was being taken seriously, citing examples of how type of task, type of team, type of job, as well as research setting, culture, and fit were recognized as moderator and mediator variables that interacted with individual difference variables such as personality to predict work behavior and performance. Her optimism may have been somewhat premature. Certainly, the controversy about the usefulness of self-report personality measures suggests that some critics who focus on the low validities of personality variables may have not accepted the fact that situational characteristics exert important moderator effects that make some validities higher and some lower across situational contexts (see Tett & Christiansen, 2007).

According to the Attraction–Selection–Attrition model (Schneider, Goldstein, & Smith, 1995), we know that individuals seek out and self-select into different jobs and work situations. For instance, we have evidence that employees perceive achievement-based or relationship-based situations in a manner consistent with their standing on measured personality traits (Westerman & Simmons, 2007). But going above and beyond these effects, some situational contexts may result in mean performance differences across individuals but do not change personality–performance correlations (i.e., situations exert main effects), and other situational contexts result in mean differences and also influence correlations between personality and performance (i.e., there are interactions). We seem to prefer simple and straightforward analyses with very few situational moderators. Yet, “the situation” is such a large and looming entity that one would be foolish to ignore its major features. For example, the relationship between cognitive variables and job performance is clearly moderated by the situational factor



of job complexity: **The relationship between cognitive ability and job performance tends to be lower for work characterized as low in complexity compared to high in complexity**, corrected  $r = .23$  for unskilled jobs versus  $.56$  for high-complexity jobs (Hunter, 1980). The relationship tends to remain constant ( $r = .56$ ), however, across job complexity when the dependent variable is performance in training programs (Schmidt & Hunter, 1998).

Regarding personality, we have evidence that the choice of personality variables based on a job analysis tends to result in higher validities than do validation studies that are exploratory in nature (Tett, Jackson, Rothstein, & Reddon, 1999). We also know that correlations between personality constructs and job-related constructs tend to be higher when item content is contextualized, that is, when the work situation is incorporated into the item (Schmit, Ryan, Stierwalt, & Powell, 1995).

The “strength” of a situation has a predictable effect on personality–performance validities. **Situational strength is the extent the situation signals expectations about what is appropriate behavior** (Mischel, 1977). In strong organizational situations, employee work behavior may be predictable without any knowledge of individual differences (e.g., highly scripted work on the assembly line), but in weak situations (e.g., individuals assemble cars as members of a team, with no assigned roles or timelines), individual behavior is more discretionary, and personality would therefore be more likely to predict work behavior. Evidence supports this notion. **In high-autonomy jobs (i.e., weak situations), validities are higher for conscientiousness, extraversion, and agreeableness than in low-autonomy jobs** (Barrick & Mount, 1993). More generally, the Big Five factors of personality show higher validities when predicting supervisory ratings of task and contextual performance in weak performance situations (when employees with similar jobs disagree about which tasks are most important) than in strong performance situations (when employees generally agree; Beaty, Cleveland, & Murphy, 2001). This

finding is in line with the research finding that personality may be more predictive under typical performance conditions, because performance is more varied across individuals, than under maximum performance conditions where all employees tend to be highly motivated, such as when they know their performance is being monitored (Klehe & Anderson, 2007; Sackett, Zedeck, & Fogli, 1988).

We call for I–O psychologists to continue work on existing situation taxonomies, such as those described by Saucier, Bel-Bahar, and Fernandez (2007), to refine and build on those taxonomies and to develop ways to incorporate them into more sophisticated moderator analyses. **We have the well-developed analytical tool of hierarchical linear modeling (HLM) to accommodate such a taxonomy, in that situational characteristics can be entered as variables that predict group differences (means and correlations)**. Furthermore, an appropriate taxonomy of personality variables should be integrated with an appropriate taxonomy of situations. We agree with Tett and Christiansen (2007) that a theory of “trait activation” (Tett & Burnett, 2003) would highlight where and how personality variables are most likely to influence behavior and performance.

#### **4. Does the combination of cognitive ability and personality measures increase the accuracy of our predictions of work performance and other work-related criteria?**

We have already discussed how there is room for improving the measurement of personality variables and job-related criteria and thus for increasing the validity of these relationships. That said, the accumulated wisdom regarding the incremental validity of personality over cognitive ability is clear:

1. **Personality variables show meaningful correlations with a wide array of work-relevant criteria** (see the previous “A Bit of History—How We Got Here” section). Even with self-report measures of personality, the validity of integrity tests is high for jobs low in complexity,  $r = .41$  (Ones et al., 1993).

2. Personality variables correlate at low levels with cognitive ability (see Ackerman & Heggstad, 1997; McHenry et al., 1990; and Ones, Viswesvaran, & Reiss, 1996, for estimates of these correlations), affording the possibility for incremental validity.
3. Personality variables do show incremental validity, thereby increasing the overall accuracy of our predictions of job performance and other job-related criteria.

Project A with its sample size of approximately 4,000 provided clear evidence of the increase in validity when both cognitive ability and personality measures are included in the same predictor battery (McHenry et al., 1990). Again, using Project A data, a model of the determinants of work performance that included personality variables accounted for more than twice the variance in supervisory performance ratings than did Hunter's (1983) model that included only cognitive ability, job knowledge, and performance (Borman, White, Pulakos, & Oppler, 1991).

We can also estimate the validity of such composites without conducting actual full-scale validity studies. If the correlations between predictors and criterion-related validities are known, the validity of the composite can be estimated (see Ghiselli, Campbell, & Zedeck, 1981; Trattner, 1982).

#### **5. If we were to better differentiate personality and other organizationally relevant constructs from the way these constructs are measured, what benefits might accrue?**

We are an applied field, and information about the usefulness of our interventions, such as personnel selection systems, is critically important. We have generated much useful knowledge about those interventions, but as a field, we have tended to confound constructs with measurement methods. In other words, we need to maintain the distinction between the psychological constructs we seek to measure and the measurement itself. **Our ability to generalize personality findings is limited to the extent they are bound to specific modes of measurement.** One of the best summaries of the

validities found in our personnel selection literature, **Schmidt and Hunter (1998)**, suffers from this problem. They summarized the literature as they found it often lacking in precision about the **differences between constructs and measurement methods** (e.g., cognitive ability, interviews, conscientiousness, assessment centers, and biographical inventories). Similarly, meta-analyses summarizing the validity of the interview have shown important results but often confound measurement method with the individual difference constructs it intends to measure. Validities tend to vary depending upon (a) **the amount of structure involved in gathering as well as evaluating interview information**, (b) **situational versus job-related versus psychological nature of the interview**, and (c) **individual versus board-conducted interview** (Huffcutt & Arthur, 1994; McDaniel, Whetzel, Schmidt, & Maurer, 1994). Interviews are likely moderated by the type of constructs measured as well. That information is difficult to determine especially in an unstructured interview, but it does not remove the confound between method and construct being assessed.

Most, if not all, of the meta-analyses of validities of personality variables for predicting work-related criteria heavily rely on self-report measures of personality. These summaries are helpful in that they provide summaries of an important and very common way of measuring personality, and self-report personality measures do predict meaningful work-related outcomes. However, without other methods of assessment, self-report and construct measurement is confounded. Additionally, we know that others' ratings of a target person's personality can result in higher correlations between personality and work-related criteria (Mount, Barrick, & Strauss, 1994; Nilsen, 1995).

In short, we need to remember the time-tested insights of our field concerning the process of construct validation that incorporate multitrait, multimethod evidence (D. T. Campbell & Fiske, 1959). At the same time we seek to accumulate stable and disambiguated information about constructs and how they are measured, many I-O

practitioners involved in personnel selection face the challenge of predicting performance in today's world of work where tasks, teams, and work contexts are constantly changing. New models for validation might involve databases that synthesize validities for ever-changing combinations of work activities and responsibilities. Even in a changing world of work, organizations need to decide who among their job applicants will add value, and they want to base those decisions on evidence that predicted performance levels are relatively accurate. Of course, this requires organizations to work with their I–O consultants to inform them what criterion constructs are important. In turn, the I–O consultants should ideally be able to conduct a work and situation analysis to identify the relevant KSAO constructs, refer to our empirical literature (both meta-analyses and primary studies), and synthesize the validity of the prediction equation to inform the organization of the utility of a recommended selection system.

#### **6. How can we reframe the “faking issue” on personality tests in more productive ways?**

Whether it has been considered a holy grail worth pursuing or a windmill at which to tilt, the topic of applicant faking in personality measures has been pursued by organizational researchers with something of a vengeance. This is not to say that research and practice has not been concerned about dissimulation on other selection measures; such research has been conducted on situational judgment tests (e.g., Cullen, Sackett, & Lievens, 2006; H. Peeters & Lievens, 2005), biographical data measures (e.g., Kluger & Collela, 1993; Schmitt & Kuncce, 2002), employment interviews (e.g., Delery & Kacmar, 1998; Ellis, West, Ryan, & DeShon, 2002; Levashina & Campion, 2006), the assessment center (e.g., McFarland, Yun, Harold, Viera, & Moore, 2005), and the ubiquitous job application blank (e.g., Wood, Schmidke, & Decker, 2007). But it is safe to say that the lion's share of research on faking selection measures in the past two decades has been directed mostly toward personality tests.

We view these longstanding concerns about faking in personality tests as legitimate in concept but suggest that the empirical research on faking could benefit from a stronger theoretical perspective and language that is focused on a framework that considers several distinct dimensions at the same time, such as:

- *test-taking mode* (e.g., paper-and-pencil, computerized, web-based)
- *test-taking settings* (e.g., timed/untimed, proctored/unproctored),
- *test-taking instructions* (e.g., to answer honestly, to answer quickly, that responses are verifiable),
- *test format* (e.g., ipsative/normative, long form/short form),
- *test-taker individual differences* (e.g., general and specific test experience, test anxiety, reading ability, and impression management),
- *external motivating factors* (e.g., tight or loose labor market, job salary and benefits), and
- *test-taking outcomes* (e.g., differences in descriptive statistics, factorial structure, and criterion-related validities).

McFarland and Ryan (2006) recently embedded some of these dimensions of applicant faking within the Theory of Planned Behavior. Tourangeau and Yan (2007) cite some of the dimensions above when explaining response distortion in items that are particularly sensitive in nature. By taking a more systematic approach to faking that keeps the practical goal of generalizability to applicant settings in mind, we are suggesting that researchers abandon the hunt for unobserved “true” scores or elusive suppressors as keys that will unlock the treasure chest of higher criterion-related validities. Many hunts of this nature have been dismally unsuccessful. In fact, simulations indicate that even going to the extreme of flagging and removing individuals with high scores on an indicator of faking or social desirability leads to a minimal increase in predicted mean performance at best (Schmitt & Oswald, 2006). This means that partialling

social desirability out of personality scores would be even less effective. **Instead of flagging extreme scores, if you assume that faking is uncorrelated with the personality constructs measured, then greater faking would in fact predict lower job performance outcomes** (Komar, Brown, Komar, & Robie, 2008); however, these assumptions are unknowable.

In real-life employment settings, relationships between substantive personality scales and job performance do not appear to be moderated or suppressed by social desirability at least as measured by scores on social desirability scales (Hough, 1998a; Hough & Ones, 2001; Hough et al., 1990; Ones et al., 1996). Studies that use instructed faking conditions can push average test scores around, particularly for “fake bad” but also for “fake good” (Hough et al., 1990; Viswesvaran & Ones, 1999); however, such studies tend to be conducted in the lab. More informative studies depart from the lab study with its college samples, and they examine the dimensions above in field settings. Ellingson, Sackett, and Hough (1999) find that job applicants with high social desirability scores still provide differentiating information across multiple personality traits. Other studies have examined applicant retest effects on personality scores after being rejected in previous application attempts. Some of those studies find negligible effects (Hogan, Barrett, & Hogan, 2007), whereas others have found quite significant effects (Young, 2003). Further research investigating these differences is warranted.

Keeping in mind the suggestions for productive avenues for future research, what are some conclusions we can make about faking based on past literature?

- Social Desirability scales, aka Unlikely Virtues scales, detect intentional distortion (Hough et al., 1990).
- Respondents can, when instructed to do so, distort their responses significantly in a positive direction, even more in a negative direction (Hough, et al., 1990; Viswesvaran & Ones, 1999).

- Social desirability response bias in real-life employment settings is not as large as that produced in directed faking studies (e.g., Dunnette, McCartney, Carlson, & Kirchner, 1962; Hansen & McLellan, 1997; Heron, 1956; Hough, 1998a; Hough et al., 1990; Kirchner, 1962; Michaelis & Eysenck, 1971; Orpen, 1971; Ryan & Sackett, 1987; Schwab & Packard, 1973; Trent, 1987; versus Rosse, Stecher, Levin, & Miller, 1998).
- In real-life employment settings, relationships between substantive personality scales and job performance do not appear to be moderated or suppressed by social desirability at least as measured by scores on social desirability scales (Hough, 1998a; Hough & Ones, 2001; Hough et al., 1990; Ones et al., 1996).
- Corrections to substantive personality scales based on response bias scale scores do not affect relationships between substantive personality scales and job performance measures (Hough, 1998a; Schmitt & Oswald, 2006).

In “bottom line” terms, hiring organizations seek to achieve the highest levels of their objectives (e.g., high performance, low turnover) in those applicants they select and train and in whom they invest time and money. Such a desire first assumes that organizations have a clear idea of how to conceptualize and operationalize their objectives (e.g., how to develop, administer, and weight multidimensional measures of job performance). To the extent that effective criterion measurement is not in place, and to the extent we cannot determine the type of test score faking that would lead to harming the organization (not hiring more effective individuals) and to the qualified applicant (being displaced by applicants who fake more), the effect of faking on criterion-related validity becomes a more difficult question to answer. We face these challenges in today’s research.

## 7. How do groups protected under U.S. Civil Rights Acts score on personality



**variables? Do personality tests have adverse impact against protected groups? Can personality tests be used in combination with cognitive ability tests to increase validity and reduce adverse impact?**

Most organizations around the world are interested in fair employment practices and employing a work force that is representative of their communities. However, when different groups score on average differently, differential hiring rates (disparate or adverse impact) can result. In the United States, organizations whose employment practices result in disparate impact against members of protected classes are at risk of charges of alleged discrimination and possible lawsuits. An often-accepted indicator of adverse impact is a selection ratio that is less than four fifth that of the majority group. An important determinant of adverse impact is a mean score difference between a protected group and the majority group on components or composites of the personnel selection measures.

Measures of cognitive ability, included in many personnel selection test batteries, often produce fairly substantial mean score differences, thereby contributing to differential hiring rates for Blacks, Hispanics, and Whites. On the other hand, adding measures of personality to a selection test battery can reduce adverse impact (J. P. Campbell, 1996). However, research, simulation results, and a mathematical analysis have all shown that countering the effects of a measure with a large mean difference between subgroups (e.g., the typical 1 SD mean Black–White difference on cognitive ability tests) requires grossly overweighting other measures showing little mean difference (e.g., DeCorte, Lievens, & Sackett, 2007; Potosky, Bobko, & Roth, 2005; Sackett & Ellingson, 1997). For instance, even though a personality measure may have a smaller subgroup mean score difference than a cognitive ability measure, if the two measures are essentially uncorrelated with each other and the majority group scores higher on both measures, then adverse impact against the protected group may actually increase when the two measures

are combined. With this general knowledge in hand, it is then important to know which measures show larger mean differences by subgroups of interest and which show smaller differences.

Hough et al. (2001) gathered and summarized standardized mean subgroup (ethnic/cultural, age, and gender) differences (*d* values) on personality, cognitive, and physical ability constructs at both broadly and more narrowly defined construct levels. They found:

. . . some surprising results. Research clearly indicates that the setting, the sample, the construct and the level of construct specificity can all, either individually or in combination, moderate the magnitude of differences between groups. Employers using tests in employment settings need to assess accurately the requirements of work. When the exact nature of the work is specified, the appropriate predictors may or may not have adverse impact against some groups (p. 152).

For example, they found very little difference between Blacks and Whites for broad-level constructs such as FFM conscientiousness and extraversion; nonetheless, at the facet level, some subgroup mean differences were as large as .2 or .3 standard deviations, with Blacks scoring higher or lower than Whites depending upon the facet. Other researchers also provide information about these important subgroup differences as well. Foldes, Duehr, and Ones (in press) meta-analyzed subgroup differences in personality variables for five U.S. racial groups for FFM variables and their facets. Else-Quest, Hyde, Goldsmith, and Van Hulle (2006) meta-analyzed gender differences in personality, and Van Vugt, De Cremer, and Janssen (2007) summarized gender differences in cooperation and competition.

These findings provide important general guidelines for personnel selection practitioners: Use job analysis information to guide the selection of tests that measure the specific needed abilities and characteristics, realizing that subgroup differences will vary



depending upon the predictor construct and its specificity, and that will affect the hiring rate of protected subgroups. Selection batteries that include measures of cognitive ability almost never satisfy the four-fifths rule, whereas selection batteries that include only measures of personality typically (but not always) satisfy the four-fifths rule (Bobko, Roth, & Potosky, 1999; Hough & Oswald, 2000; Ryan, Ployhart, & Friedel, 1998; Schmitt, Rogers, Chan, Sheppard, & Jennings, 1997). Group mean differences and the relative weights for all the cognitive and personality variables contributing to a predictor composite should be carefully examined before making any assumptions about the composite. Campion et al. (2001) addressed the dilemma of the twin, but often competing, goals of achieving accurate prediction (criterion-related validity) while also achieving similar selection rates for subgroups (reduced adverse impact). Using linear programming methods, De Corte et al. (2007) proposed a procedure for forming a weighted composite that reduces adverse impact as much as possible, given a specified level of validity. We suggest examining this procedure to understand the sensitivity of predictor weights on adverse impact and validity outcomes. In particular, regression weights are designed to maximize validity in a sample and will therefore capitalize on chance to some extent. A different set of predictor weights that is robust across samples would show more consistent validities that do not capitalize on chance, and they may also demonstrate reduced adverse impact. Job-relevant predictor composites often contain cognitive ability measures with their likely adverse impact, but new statistical methods can help us understand these issues in a more precise manner.

## Summary

In the 1980s, Project A reoriented our field's thinking about both the predictor and the criterion space, focusing on constructs and multifaceted models of performance. Personality constructs were shown to relate to performance constructs in meaningful ways.

Since the 1990s, there has been a near-literal explosion of interest and research involving personality constructs and multidimensional models of performance that continues to this day. It has become clear that the complexity of the nomological nets of personality constructs is enormous. In comparison with cognitive ability, which is hierarchical and often driven by *g* in its validity, we have only begun to scratch the surface. Our theories and models of personality-based determinants of behavior and performance need to catch up with this inherent complexity. We need better models, better interventions, and more evidence-based practice as they relate to behavior and performance in organizational settings.

Our goal was to provide a current perspective on personality testing and its use in organizational research and practice by offering a series of questions about personality testing to spur constructive discussion and future research and practice in I-O psychology.

## References

- Ackerman, P. L., & Heggstad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, 121, 219–245.
- Ashton, M. C. (1998). Personality and job performance: The importance of narrow traits. *Journal of Organizational Behavior*, 19, 289–303.
- Barrick, M. R., & Mount, M. K. (1993). Autonomy as a moderator of the relationships between the big five personality dimensions and job performance. *Journal of Applied Psychology*, 78, 111–118.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1–26.
- Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, 9, 9–30.
- Barrick, M. R., Mount, M. K., & Strauss, J. P. (1993). Conscientiousness and performance of sales representatives: Test of the mediating effects of goal setting. *Journal of Applied Psychology*, 78, 715–722.
- Barrick, M. R., Stewart, G. L., & Piotrowski, M. (2002). Personality and job performance: Test of the mediating effects of motivation among sales representatives. *Journal of Applied Psychology*, 87, 43–51.
- Bartram, D. (2005). The Great Eight competencies: A criterion-centric approach to validation. *Journal of Applied Psychology*, 90, 1185–1203.
- Baumeister, R. F., Gailliot, M., DeWall, C. N., & Oaten, M. (2006). Self-regulation and personality: How

- interventions increase regulatory success, and how depletion moderates the effect of traits on behavior. *Journal of Personality*, 74, 1773–1801.
- Beatty, J. C., Cleveland, J. N., & Murphy, K. R. (2001). The relation between personality and contextual performance in “strong” versus “weak” situations. *Human Performance*, 14, 125–148.
- Berry, C. M., Ones, D. S., & Sackett, P. R. (2007). Interpersonal deviance, organizational deviance, and their common correlates: A review and meta-analysis. *Journal of Applied Psychology*, 92, 410–424.
- Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, 117, 187–215.
- 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, 1–31.
- Bogg, T., & Roberts, B. W. (2004). Conscientiousness and health behaviors: A meta-analysis of the leading behavioral contributors to mortality. *Psychological Bulletin*, 130, 887–919.
- Bono, J. E., & Judge, T. A. (2004). Personality and transformational and transactional leadership: A meta-analysis. *Journal of Applied Psychology*, 89, 901–910.
- Borman, W. C. (2004). The concept of organizational citizenship. *Current Directions in Psychological Science*, 6, 238–241.
- Borman, W. C., Penner, L. A., Allen, T. D., & Motowidlo, S. J. (2001). Personality predictors of citizenship performance. *International Journal of Selection and Assessment*, 9, 52–69.
- Borman, W. C., White, L. A., Pulakos, E. D., & Oppler, S. H. (1991). Models of supervisory job performance ratings. *Journal of Applied Psychology*, 76, 863–872.
- Brunswik, E. (1952). The conceptual framework of psychology. Chicago: University of Chicago Press.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validity in the multi-trait multi-method matrix. *Psychological Bulletin*, 56, 81–105.
- Campbell, J. P. (1990). The role of theory in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 39–73). Palo Alto, CA: Consulting Psychologists Press.
- Campbell, J. P. (1996). Group differences and personnel decisions: Validity, fairness, and affirmative action. *Journal of Vocational Behavior*, 49, 122–158.
- Campbell, J. P., & Knapp, D. (2001). *Project A: Exploring the limits of performance improvement through personnel selection and classification*. Hillsdale, NJ: Erlbaum.
- Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 35–70). San Francisco: Jossey Bass.
- Campion, M. A., Outtz, J. L., Zedeck, S., Schmidt, F. L., Kehoe, J. F., Murphy, K. R., et al. (2001). The controversy over score banding in personnel selection: Answers to 10 key questions. *Personnel Psychology*, 54, 149–185.
- Colquitt, J. A., LePine, J. A., & Noe, R. A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology*, 85, 678–707.
- Cooper-Hakim, A., & Viswesvaran, C. (2002). A meta-analytic review of the MacAndrew Alcoholism scale. *Educational and Psychological Measurement*, 62, 818–829.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302.
- Cullen, M. J., Sackett, P. R., & Lievens, F. (2006). Threats to the operational use of situational judgment tests in the college admissions process. *International Journal of Selection and Assessment*, 14, 142–155.
- Dalal, R. (2005). A meta-analysis of the relationship between organizational citizenship behavior and counterproductive work behavior. *Journal of Applied Psychology*, 90, 1241–1255.
- De Corte, W., Lievens, F., & Sackett, P. R. (2007). Combining predictors to achieve optimal trade-offs between selection quality and adverse impact. *Journal of Applied Psychology*, 92, 1380–1393.
- Delery, J. E., & Kacmar, K. M. (1998). The influence of applicant and interviewer characteristics on the use of impression management. *Journal of Applied Social Psychology*, 28, 1649–1669.
- Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A meta-analytic investigation of Conscientiousness in the prediction of job performance: Examining the intercorrelations and the incremental validity of narrow traits. *Journal of Applied Psychology*, 91, 40–57.
- Dunnette, M. D. (1963). A note on the criterion. *Journal of Applied Psychology*, 47, 251–254.
- Dunnette, M. D., McCartney, J., Carlson, H. C., & Kirchner, W. K. (1962). A study of faking behavior on a forced-choice self-description checklist. *Personnel Psychology*, 15, 13–24.
- Ellingson, J. E., Sackett, P. R., & Hough, L. M. (1999). Social desirability correction in personality measurement: Issues of applicant comparison and construct validity. *Journal of Applied Psychology*, 84, 155–166.
- Ellis, A. P. J., West, B. J., Ryan, A. M., & DeShon, R. P. (2002). The use of impression management tactics in structured interviews: A function of question type? *Journal of Applied Psychology*, 87, 1200–1208.
- Else-Quest, N. M., Hyde, J. S., Goldsmith, H. H., & Van Hulle, C. A. (2006). Gender differences in temperament: A meta-analysis. *Psychological Bulletin*, 132, 33–72.
- Erez, A., & Judge, T. A. (2001). Relationship of core self-evaluations to goal setting, motivation, and performance. *Journal of Applied Psychology*, 86, 1270–1279.
- Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review*, 2, 290–309.
- Fleishman, E. A., & Quaintance, M. K. (1984). Taxonomies of human performance: *The description of human tasks*. Orlando, FL: Academic Press.
- Foldes, H., Duehr, E. E., & Ones, D. S. (in press). Group differences in personality: Meta-analyses comparing five U.S. racial groups. *Personnel Psychology*.
- Ghiselli, E. E. (1966). *The validity of occupational aptitude tests*. New York: John Wiley & Sons.
- Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences*. San Francisco: W.H. Freeman.
- Guion, R. M. (1961). Criterion measurement and personnel judgments. *Personnel Psychology*, 14, 141–149.

- Guion, R. M., & Gottier, R. F. (1965). Validity of personality measures in personnel selection. *Personnel Psychology, 18*, 135–164.
- Hansen, T. L., & McLellan, R. A. (1997, April). Social desirability and item content. In G. J. Lautenschlager (Chair), *Faking on non-cognitive measures: The extent, impact, and identification of dissimulation*. Symposium conducted at the 12th Annual Conference of the Society of Industrial and Organizational Psychology, St. Louis, MO.
- Hatrup, K., & Jackson, S. E. (1996). Learning about individual differences by taking situations seriously. In K. R. Murphy (Ed.), *Individual differences and behavior in organizations* (pp. 507–541). San Francisco: Jossey-Bass.
- Heron, A. (1956). The effects of real-life motivation on questionnaire response. *Journal of Applied Psychology, 40*, 65–68.
- Hogan, J., Barrett, P., & Hogan, R. (2007). Personality measurement, faking, and employment selection. *Journal of Applied Psychology, 92*, 1270–1285.
- Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. *Journal of Applied Psychology, 88*, 100–112.
- Hogan, J., & Ones, D. S. (1997). Conscientiousness and integrity at work. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 513–541). San Diego: Academic Press.
- Hollenbeck, J. R. (1998). *Personnel Psychology's* citation leading articles: The first five decades [Introduction]. *Personnel Psychology, 51*, Editorial.
- Hough, L. M. (1992). The “Big Five” personality variables—Construct confusion: Description versus prediction. *Human Performance, 5*, 139–155.
- Hough, L. M. (1997). The millennium for personality psychology: New horizons or good old daze. *Applied Psychology: An International Review, 47*, 233–261.
- Hough, L. M. (1998a). Effects of intentional distortion in personality measurement and evaluation of suggested palliatives. *Human Performance, 11*, 209–244.
- Hough, L. M. (1998b). Personality at work: Issues and evidence. In M. Hakel (Ed.), *Beyond multiple choice: Evaluating alternatives to traditional testing for selection* (pp. 131–166). Hillsdale, NJ: Lawrence Erlbaum.
- Hough, L. M. (2001). I/O owes its advances to personality. In B. Roberts & R. T. Hogan (Eds.), *Personality psychology in the workplace* (pp. 19–44). Washington, DC: American Psychological Association.
- Hough, L. M. (2003). Emerging trends and needs in personality research and practice: Beyond main effects. In M. Barrick & A. Ryan (Eds.), *Personality and work* (pp. 289–325). New York: Wiley & Sons.
- Hough, L. M., & Dilchert, S. (2007, October). Inventors, innovators, and their leaders: Selecting for Conscientiousness will keep you “inside the box.” Paper presented at SIOP's 3rd Leading Edge Consortium: *Enabling Innovation in Organizations*, Kansas City, MO.
- Hough, L. M., Eaton, N. L., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities [Monograph]. *Journal of Applied Psychology, 75*, 581–595.
- Hough, L. M., & Furnham, A. (2003). Importance and use of personality variables in work settings. In I. B. Weiner (Ed-in-Chief) & W. Borman, D. Ilgen, & R. Klimoski (Vol. Eds.), *Comprehensive handbook of psychology: Industrial organizational psychology*. (Vol. 12, pp. 131–169). New York: Wiley & Sons.
- Hough, L. M., & Ones, D. S. (2001). The structure, measurement, validity, and use of personality variables in industrial, work, and organizational psychology. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work and organizational psychology* (Vol. 1, pp. 233–377). London: Sage.
- Hough, L. M., Ones, D. S., & Viswesvaran, C. (1998, April). Personality correlates of managerial performance constructs. In R. Page (Chair), *Personality determinants of managerial potential, performance, progression and ascendancy*. Symposium conducted at 13th Annual Conference of the Society for Industrial and Organizational Psychology, Dallas, TX.
- Hough, L. M., & Oswald, F. L. (2000). Personnel selection: Looking toward the future—Remembering the past. *Annual Review of Psychology, 51*, 631–664.
- Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection, and amelioration of adverse impact in personnel selection procedures: Issues, evidence, and lessons learned. *International Journal of Selection and Assessment, 9*, 152–194.
- Hough, L. M., & Schneider, R. J. (1996). Personality traits, taxonomies, and applications in organizations. In K. Murphy (Ed.), *Individual differences and behavior in organizations* (pp. 31–88). San Francisco: Jossey-Bass.
- Huffcutt, A. I., & Arthur, W., Jr. (1994). Hunter and Hunter (1984) revisited: Interview validity for entry-level jobs. *Journal of Applied Psychology, 79*, 184–190.
- Hunter, J. E. (1980). *Validity generalization for 12,000 jobs: An application of synthetic validity and validity generalization to the General Aptitude Test Battery (GATB)*. Washington, DC: U.S. Department of Labor, Employment Service.
- Hunter, J. E. (1983). A causal analysis of cognitive ability, job knowledge, job performance, and supervisor ratings. In F. Landy, S. Zedeck, & J. Cleveland (Eds.), *Performance measurement and theory* (pp. 257–266). Mahwah, NJ: Erlbaum.
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology, 85*, 869–879.
- Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology, 87*, 765–780.
- Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2003). The core self-evaluation scale: Development of a measure. *Personnel Psychology, 56*, 303–331.
- Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology, 87*, 530–541.
- Judge, T. A., & Ilies, R. (2002). Relationship of personality to performance motivation: A meta-analytic review. *Journal of Applied Psychology, 87*, 797–807.
- Kamp, J. D., & Hough, L. M. (1986). Utility of personality assessment: A review and integration of the

- literature. In L. M. Hough (Ed.), *Utility of temperament, biodata, and interest assessment for predicting job performance: A review and integration of the literature* (ARI Research Note No. 88-02, pp. 1–90). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Kanfer, R., Ackerman, P. L., Murtha, T., & Goff, M. (1995). Personality and intelligence in industrial and organizational psychology. In D. H. Saklofske & M. Zeidner (Eds.), *International handbook of personality and intelligence* (pp. 577–602). New York: Plenum Publishing.
- Kanfer, R., & Heggestad, E. D. (1999). Individual differences in motivation: Traits and self-regulatory skills. In P. L. Ackerman & P. C. Kyllonen (Eds.), *Learning and individual differences: Process, trait, and content determinants* (pp. 293–313). Washington, DC: American Psychological Association.
- Kirchner, W. K. (1962). "Real-life" faking on the Edwards Personal Preference Schedule by sales applicants. *Journal of Applied Psychology*, 46, 128–130.
- Klehe, U. C., & Anderson, N. (2007). Working hard and working smart: Motivation and ability during typical and maximum performance. *Journal of Applied Psychology*, 92, 978–992.
- Klein, H. J., & Lee, S. (2006). The effects of personality on learning: The mediating role of goal setting. *Human Performance*, 19, 43–66.
- Kluger, A. N., & Collela, A. (1993). Beyond the mean bias: Effects of warning against faking on biodata item variances. *Personnel Psychology*, 46, 763–780.
- Kohlberg, L., Levine, C., & Hewer, A. (1994). Moral stages: A current formulation and a response to critics. In B. Puka (Ed.), *New research in moral development (Moral development: A compendium)* (Vol. 5, pp. 126–188). New York: Garland Publishing.
- Komar, S., Brown, D. G., Komar, J. A., & Robie, C. (2008). Faking and the validity of conscientiousness: A Monte Carlo investigation. *Journal of Applied Psychology*, 93, 140–154.
- Lee, F. K., Sheldon, K. M., & Turban, D. B. (2003). Personality and the goal-striving process: The influence of achievement goal patterns, goal level, and mental focus on performance and enjoyment. *Journal of Applied Psychology*, 88, 256–265.
- Lee, K., Ashton, M. C., & deVries, R. E. (2005). Predicting workplace delinquency and integrity with the HEXACO and five-factor models of personality structure. *Human Performance*, 18, 179–197.
- LePine, J. A., Erez, A., & Johnson, D. E. (2002). The nature and dimensionality of organizational citizenship behavior: A critical review and meta-analysis. *Journal of Applied Psychology*, 87, 52–65.
- Levashina, J., & Campion, M. A. (2006). A model of faking likelihood in the employment interview. *International Journal of Selection and Assessment*, 14, 299–316.
- Loevinger, J. (1966). The meaning and measurement of ego development. *American Psychologist*, 21, 195–206.
- Loevinger, J. (1976). *Ego development: Conceptions and theories*. San Francisco: Jossey-Bass.
- Loevinger, J. (1994). In search of grand theory. *Psychological Inquiry*, 5, 142–144.
- McAdams, D. P. (1992). The five-factor model in personality: A critical appraisal. *Journal of Personality*, 60, 329–361.
- McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, 79, 599–616.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.
- McFarland, L., & Ryan, A. M. (2006). Toward an integrated model of applicant faking behavior. *Journal of Applied Social Psychology*, 36, 979–1016.
- McFarland, L., Yun, G., Harold, C. M., Viera, R., Jr., & Moore, L. (2005). An examination of impression management use and effectiveness across assessment center exercises: The role of competency demands. *Personnel Psychology*, 58, 949–980.
- 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.
- Merston, B., & Gorsuch, R. L. (1988). Number of factors in personality sphere: Does increase in factors increase predictability of real life criteria? *Journal of Personality and Social Psychology*, 55, 675–680.
- Michaelis, W., & Eysenck, H. J. (1971). The determination of personality inventory factor pattern and intercorrelations by changes in real-life motivation. *Journal of Genetic Psychology*, 118, 223–234.
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
- Mischel, W. (1977). The interaction of person and situation. In D. Magnusson & N. S. Endler (Eds.), *Personality at the crossroads: Current issues in interactional psychology* (pp. 333–352). Hillsdale, NJ: Erlbaum.
- Mol, S. T., Born, M. P. H., Willemsen, M. E., & Van Der Molen, H. T. (2005). Predicting expatriate job performance for selection purposes: A quantitative review. *Journal of Cross-Cultural Psychology*, 36, 590–620.
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007a). Are we getting fooled again? Coming to terms with limitations in the use of personality tests for personnel selection. *Personnel Psychology*, 60, 1029–1049.
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007b). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology*, 60, 683–729.
- Mount, M. K., Barrick, M. R., & Strauss, J. P. (1994). Validity of observer ratings of the Big Five personality factors. *Journal of Applied Psychology*, 79, 272–280.
- Nilsen, D. (1995). *Investigation of the relationship between personality and leadership performance*. Unpublished doctoral dissertation, Minneapolis, MN: University of Minnesota.
- Ng, T. W. H., Eby, L. T., Sorensen, K. L., & Feldman, D. C. (2005). Predictors of objective and subjective career success: A meta-analysis. *Personnel Psychology*, 58, 367–408.
- Ones, D. S., Dilchert, S., Viswesvaran, C., & Judge, T. A. (2007). In support of personality assessment in organizational settings. *Personnel Psychology*, 60, 995–1027.

- Ones, D. S., Viswesvaran, C., & Dilchert, S. (2005). Personality at work: Raising awareness and correcting misconceptions. *Human Performance*, 18, 389–404.
- Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology*, 81, 660–679.
- Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology*, 78, 679–703.
- Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (2003). Personality and absenteeism: A meta-analysis of integrity tests. *European Journal of Personality*, 17(Suppl.), S19–S38.
- Ong, A. D., Bergeman, C. S., Bisconti, T. L., & Wallace, K. A. (2006). Psychological resilience, positive emotions, and successful adaptation to stress in later life. *Journal of Personality and Social Psychology*, 91, 730–749.
- Organ, D. W., & Ryan, K. (1995). A meta-analytic review of attitudinal and dispositional predictors of organizational citizenship behavior. *Personnel Psychology*, 48, 775–802.
- Orpen, C. (1971). Fakability of the Edwards Personal Preference Schedule in personnel selection. *Personnel Psychology*, 24, 1–4.
- Paunonen, S. V. (1998). Hierarchical organization of personality and prediction of behavior. *Journal of Personality and Social Psychology*, 74, 538–556.
- Paunonen, S. V., & Ashton, M. C. (2001). Big Five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology*, 81, 524–539.
- Peeters, H., & Lievens, F. (2005). Situational judgment tests and their predictiveness of college students' success: The influence of faking. *Educational and Psychological Measurement*, 65, 70–89.
- Peeters, A. G., Van Tuijl, H. F. J. M., Rutte, C. G., & Reymen, I. M. M. J. (2006). Personality and team performance: A meta-analysis. *European Journal of Personality*, 20, 377–396.
- Pervin, L. A. (1994). A critical analysis of current trait theory. *Psychological Inquiry*, 5, 103–113.
- Potosky, D., Bobko, P., & Roth, P. L. (2005). Forming composites of cognitive ability and alternative measures to predict job performance and reduce adverse impact: Corrected estimates and realistic expectations. *International Journal of Selection and Assessment*, 13, 304–315.
- Roberts, B. W., Chernyshenko, O. S., Stark, S., & Goldberg, L. R. (2005). The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires. *Personnel Psychology*, 58, 103–139.
- Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, 2, 313–345.
- Rosse, J. G., Stecher, M. D., Levin, R. A., & Miller, J. L. (1998). The impact of response distortion on preemployment personality testing and hiring decisions. *Journal of Applied Psychology*, 83, 634–644.
- Rothstein, M. G., & Goffin, R. D. (2006). The use of personality measures in personnel selection: What does current research support? *Human Resource Management Review*, 16, 155–180.
- 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, 298–307.
- Ryan, A. M., & Sackett, P. R. (1987). Pre-employment honesty testing: Fakability, reactions of test takers, and company image. *Journal of Business and Psychology*, 1, 248–256.
- Sackett, P. R., Berry, C. M., Wiemann, S. A., & Lacz, R. M. (2006). Citizenship and counterproductive behavior: Clarifying relations between the two domains. *Human Performance*, 19, 441–464.
- Sackett, P. R., & DeVore, C. J. (2002). Counterproductive behaviors at work. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work and organizational psychology* (Vol. 1, pp. 145–164). Thousand Oaks, CA: Sage.
- 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.
- Sackett, P. R., Zedeck, S., & Fogli, L. (1988). Relations between measures of typical and maximum job performance. *Journal of Applied Psychology*, 73, 482–486.
- Saucier, G., Bel-Bahar, T., & Fernandez, C. (2007). What modifies the expression of personality tendencies? Defining basic domains of situation variables. *Journal of Personality*, 75, 479–503.
- Saucier, G., & Goldberg, L. R. (1998). What is beyond the Big Five? *Journal of Personality*, 66, 495–524.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implication of 85 years of research findings. *Psychological Bulletin*, 124, 262–274.
- Schmitt, M. J., Ryan, A. M., Stierwalt, S. L., & Powell, A. B. (1995). Frame-of-reference effects on personality scale scores and criterion-related validity. *Journal of Applied Psychology*, 80, 607–620.
- Schmitt, N., & Kunc, C. (2002). The effects of required elaboration of answers to biodata questions. *Personnel Psychology*, 55, 569–587.
- Schmitt, N., & Oswald, F. L. (2006). The impact of corrections for faking on the validity of noncognitive measures in selection settings. *Journal of Applied Psychology*, 91, 613–621.
- Schmitt, N., Rogers, W., Chan, D., Sheppard, L., & Jennings, D. (1997). Adverse impact and predictive efficiency of various predictor combinations. *Journal of Applied Psychology*, 82, 719–730.
- Schneider, B., Goldstein, H. W., & Smith, D. B. (1995). The ASA framework: An update. *Personnel Psychology*, 48, 747–779.
- Schneider, R. J., & Hough, L. M. (1995). Personality and industrial/organizational psychology. In C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology* (pp. 75–129). Chichester, UK: Wiley.
- Schneider, R. J., Hough, L. M., & Dunnette, M. D. (1996). Broadened by broad traits: How to sink science in five dimensions or less. *Journal of Organizational Behavior*, 17, 639–655.



- Schwab, D. P., & Packard, G. L. (1973). Response distortion on the "Gordon Personal Inventory" and the "Gordon Personal Profile" in a selection context: Some implications for predicting employee tenure. *Journal of Applied Psychology, 58*, 372-374.
- Smith, P. C. (1976). Behaviors, results, and organizational effectiveness: The problem of criteria. In M. D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 745-775). Chicago: Rand McNally College Publishing.
- Steel, P. (2007). The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological Bulletin, 133*, 65-94.
- Steel, P., Schmidt, J., & Shulz, J. (2008). Refining the relationship between personality and subjective well-being. *Psychological Bulletin, 134*, 138-161.
- Stewart, G. (1998). Trait bandwidth and stages of job performance: Assessing differential effects for conscientiousness and its subtraits. *Journal of Applied Psychology, 84*, 959-968.
- Tellegen, A. (1993). Folk concepts and psychological concepts of personality and personality disorder. *Psychological Inquiry, 4*, 122-130.
- Tett, R. P., & Burnett, D. B. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology, 88*, 500-517.
- Tett, R. P., & Christiansen, N. D. (2007). Personality tests at the crossroads: A response to Morgeson, Campion, Dipboye, Hollenbeck, Murphy, and Schmitt (2007). *Personnel Psychology, 60*, 967-993.
- Tett, R. P., Jackson, D. N., Rothstein, M., & Reddon, J. R. (1999). Meta-analysis of bi-directional relations in personality-job performance research. *Human Performance, 12*, 1-29.
- Tett, R. P., Steele, J. R., & Beauregard, R. S. (2003). Broad and narrow measures on both sides of the personality-job performance relationship. *Journal of Organizational Behavior, 24*, 335-356.
- Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin, 133*, 859-883.
- Trattner, M. H. (1982). Synthetic validity and its application to the Uniform Guidelines validation requirements. *Personnel Psychology, 35*, 383-397.
- Trent, T. (1987, August). *Armed forces adaptability screening: The problem of item response distortion*. Paper presented at the 95th Annual Convention of the American Psychological Convention, New York City.
- Van Vugt, M., De Cremer, D., & Janssen, D. P. (2007). Gender differences in cooperation and competition: The male-warrior hypothesis. *Psychological Science, 18*, 19-23.
- Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement, 59*, 197-210.
- Westerman, J. W., & Simmons, B. L. (2007). The effects of work environment on the personality-performance relationship: An exploratory study. *Journal of Managerial Issues, 19*, 288-305.
- Wiggins, J. S., & Trapnell, P. D. (1997). Personality structure: The return of the Big Five. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 737-765). San Diego: Academic.
- Wood, J. L., Schmidke, J. M., & Decker, D. L. (2007). Lying on job applications: The effects of job relevance, commission, and human resource management experience. *Journal of Business Psychology, 22*, 1-9.
- Young, M. C. (2003, June). *Effects of retesting on a new Army measure of motivational attributes: Implications for response distortion, test validity, and operational use*. Paper presented at the 27th IPMAAC Conference on Personnel Assessment, Baltimore, MD.

Copyright of Industrial & Organizational Psychology is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.