

Running Head: GENERALIZING PERSONALITY

Generalizing Personality-Based Validity Evidence
Joyce Hogan, Scott Davies, and Robert Hogan
Hogan Assessment Systems

Chapter to appear in S. M. McPhail (Ed.), *Alternative validation strategies*.

Introduction

Prior to 1977, criterion-related validity research involved testing the hypothesis that a particular predictor variable (e.g., a cognitive ability measure) covaried reliably with a particular criterion variable (e.g., performance in training). Researchers then repeated this test using different samples, predictors, and criterion measures. Not surprisingly, results from these studies often differed between locations with similar jobs, and this variability made firm generalizations difficult. More importantly, this variability challenged the scientific integrity of the entire enterprise of personnel selection.

Researchers often explained the differences in study results in terms of situational specificity, the view that the validity of a measure is specific to the contexts and jobs under study (Gatewood & Feild, 1994; Ghiselli, 1966; Ghiselli & Brown, 1955); these differences required conducting separate validation studies for each organization, job, or group of employees. Using a large database, Schmidt and Hunter (1977) presented evidence showing that the variability in validity coefficients in single-location studies is due to statistical and procedural factors (Guion, 1998, p. 368)—idiosyncratic factors that can be ignored or statistically corrected.

Schmidt and Hunter introduced meta-analysis to psychometric research; meta-analysis is a methodology for aggregating correlation coefficients from independent studies testing the same hypothesis. They argued that differences in a test's validity across studies reflect statistical artifacts (e.g., sampling deficiency) and measurement problems (e.g., predictor/criterion unreliability, range restriction) and not unique jobs or situations. Subsequent research suggests that the correlations between performance measures and cognitive ability tests (Schmidt & Hunter, 1977), biographical data inventories (Schmidt & Rothstein, 1994), personality inventories (Barrick & Mount, 1991, 2003; Hogan & Holland, 2003; Hough, 1992; Judge, Bono,

Ilies, & Gerhardt, 2002; Salgado, 1997, 1998; Tett, Jackson, & Rothstein, 1991), assessment center exercises (Arthur, Day, McNelly, & Edens, 2003; Gaugler, Rosenthal, Thornton, & Bentson, 1987), and situational judgment tests (McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001) generalize across studies.

Validity generalization (VG) evidence, when available, may be used in place of local validation studies to support the use of a selection procedure (Gatewood & Feild, 1994; Society for Industrial and Organizational Psychology, 2003). As indicated by the *Principles for the Validation and Use of Personnel Selection Procedures (Principles; Society for Industrial and Organizational Psychology, 2003)*:

At times, sufficient accumulated validity evidence is available for a selection procedure to justify its use in a new situation without conducting a local validation research study. In these instances, use of the selection procedure may be based on demonstration of the generalized validity inferences from that selection procedure, coupled with a compelling argument for its applicability to the current situation. Although neither mutually exclusive nor exhaustive, several strategies for generalizing validity evidence have been delineated: (a) transportability, (b) synthetic validity/job component validity, and (c) meta-analytic validity generalization (p. 27).

This chapter illustrates the use of personality measures in VG research. We describe a conceptual model that links personality variables to workplace behavior. We illustrate strategies for using validity evidence from previous studies to apply to new jobs. We use research from managerial jobs as examples of the VG strategy. We then discuss the limitations of these methods and considerations for VG analyses.

Personality Measurement

Most important personality characteristics can be described in terms of five broad dimensions. The Five-Factor Model (FFM; cf. De Raad & Perugini, 2002; Digman, 1990; Goldberg, 1992; John, 1990, p. 72; McCrae & Costa, 1987; Wiggins, 1996), which is based on fifty years of factor analytic research on the structure of observer ratings (cf. Norman, 1963;

Thurstone, 1934; Tupes & Christal, 1961), suggests that we think about and describe people (Goldberg, 1990) in terms of the following themes:

- I. *Surgency/Extraversion* - the degree to which a person seems outgoing and talkative.
- II. *Agreeableness* - the degree to which a person seems pleasant and rewarding to deal with.
- III. *Conscientiousness* - the degree to which a person complies with rules, norms, and standards.
- IV. *Emotional Stability* - the degree to which a person appears calm and self-accepting.
- V. *Intellect/Openness to Experience* - the degree to which a person seems creative and open-minded.

The FFM is the basis for several personality inventories constructed over the last twenty years (e.g., NEO-PI: Costa & McCrae, 1985; Hogan Personality Inventory: R. Hogan & Hogan, 1995; Personal Characteristics Inventory: Mount & Barrick, 2001). The FFM provides a useful taxonomy for classifying individual differences in social behavior. In addition, evidence suggests that all existing multidimensional personality inventories can be reconfigured in terms of these five dimensions (De Raad & Perugini, 2002; Wiggins & Pincus, 1992). Consequently, the FFM is useful for classifying personality predictors for VG research (De Raad & Perugini, 2002; Hogan & Holland, 2003).

The FFM is based on observers' descriptions of others, and those descriptions reflect reputation (Hogan, 1983; Hogan & Shelton, 1998). Reputations are based on social consensus regarding consistencies in another person's behavior, as seen in his/her interactions at work and elsewhere—reputations are ubiquitous and enduring facts of social life. Peoples' social behavior consists, at least in part, of efforts to establish, defend, or enhance their identity – i.e., a person's view of him or herself (cf. Goffman, 1958). Reputations are public, they tell us about observable

tendencies in the others' behaviors, they can be measured reliably, and they can be used to forecast future behavior (cf. Emler, 1990; Ozer & Benet-Martinez, 2006). A person's reputation is an invaluable source of information about work-related strengths and shortcomings, and it also influences the direction of careers.

Personality assessment samples self-presentational behavior – i.e., how a person portrays him or herself to others across situations. Assessment instruments allow us to aggregate these behavioral samples, assign them numbers or scores according to certain agreed-upon rules, and then use these scores to make predictions about a person's future behavior. The research described in this chapter used the Hogan Personality Inventory (HPI: R. Hogan & Hogan, 1995) to provide data for generalizing validity evidence. The HPI was the first measure of normal personality developed explicitly to assess the FFM in everyday life. The HPI was developed using samples of working adults, and its measurement goal is to predict real-world outcomes—including job performance.

The HPI is a 206-item true-false inventory that contains seven primary scales that are aligned with the FFM. The HPI splits the FFM dimension of Extraversion into Ambition and Sociability (cf. R. Hogan & Hogan, 1995, p. 11) and Intellect-Openness to Experience into Inquisitive—which reflects creativity--and Learning Approach—which reflects achievement orientation. The internal consistency and test-retest reliability of the scales are as follows: Adjustment (.89/.86), Ambition (.86/.83), Sociability (.83/.79), Interpersonal Sensitivity (.71/.80), Prudence (.78/.74), Inquisitive (.78/.83), and Learning Approach (.75/.86). Hogan and Holland (2003) summarize considerable validity evidence for predicting occupational criteria.

Why Personality Matters at Work: A Domain Model of Job Performance

To understand why the validity of personality measures generalizes across contexts, we first need to organize social behavior in terms of its underlying motivational themes. There are two such themes: (1) most people want to be liked and accepted by others (“getting along”); (2) most people want power, status, and success (“getting ahead”). Next, we need a framework for classifying performance. We approach this by classifying competencies; we believe every competency model that has been proposed can be classified in terms of four domains which we call Intrapersonal, Interpersonal, Technical, and Leadership (cf. Bartram, 2005; Hogan & Warrenfeltz, 2003). The structure of our performance model appears in Table 1. Previous atheoretical reviews of predictor-performance relationships provide are useful for estimating the consequences of using a particular class of measures for selection decisions. In principle, however, theory-driven data aggregation should lead to results that are more generalizable. The domain model provides a systematic account of the links between personality and occupational performance.

Metaconcepts and Job Performance Models

We now review this conceptual model in somewhat more detail. Socioanalytic theory (Hogan, 1983;1996) proposes two generalizations about human nature: (1) people always live (work) in groups; and (2) groups are always structured in terms of status hierarchies. These generalizations suggest two conclusions. First, people need acceptance and approval and want to avoid rejection. These needs translate into behavior designed to “get along” with other members of the group. Supervisors and peers evaluate individuals who are skilled at “getting along” as team players who are rewarding people with whom to deal. Second, people need status and control of resources, and they try to avoid losing them. At work, these needs translate into behavior designed to “get ahead” or achieve status vis á vis other members of the group.

Supervisors and peers evaluate individuals who are skilled at “getting ahead” as having initiative and getting results.

Getting along and getting ahead are traditional themes in personality psychology and are often called agency and communion (cf. Adler, 1939; Bakan, 1966; Rank, 1945; Wiggins & Trapnell, 1996). Empirical research also shows that these constructs are related to performance outcomes (cf. Barrick, Stewart, & Piotrowski, 2002). Additionally, anthropologists (e.g., Redfield, 1960, p. 345) note that societies depend on their members “getting a living and living together.” “Getting a living” concerns being successful at life tasks, and “living together” concerns maintaining group solidarity—*esprit de corps*. There are solid conceptual and empirical reasons for organizing the structure of job performance around these metaconcepts of getting along and getting ahead.

Other models of job performance also reflect these themes. Campbell, McHenry, and Wise (1990) proposed that performance in entry level jobs in the U.S. Army could be evaluated in terms of five dimensions: core proficiency, general soldier proficiency, effort and leadership, personal discipline, and physical fitness/military bearing. Campbell, McCloy, Oppler, and Sager (1993) expanded this taxonomy to include eight factors: job-specific task proficiency, non-job-specific task proficiency, written and oral communication task proficiency, demonstrating effort, maintaining personal discipline, facilitating peer and team performance, supervision/leadership, and management/administration. In Campbell’s models, “getting along” is represented by personal discipline and facilitating peer and team performance, whereas “getting ahead” is represented by the proficiency and leadership dimensions.

Borman and Motowidlo (1993) distinguish between task performance and contextual performance—which is non-task performance that is important in all jobs. Task performance

seems to correspond to getting ahead and contextual performance corresponds to getting along with others. Borman and Motowidlo (1993) then construct a five-dimensional model of contextual performance. Their non-task criterion space overlaps considerable with that described by Campbell, Gasser, and Oswald (1996), and Rotundo and Sackett (2002). In addition, Motowidlo, Borman, and Schmit (1997) describe performance in terms of episodes, and evaluate it in terms of the perspective of the observer—i.e., they organize occupational episodes around the agendas and roles that people use to get along and get ahead.

Hunt (1996) proposed a nine-dimensional model of entry-level job performance based on factor analyses of performance ratings. His model highlights technical proficiency, but also emphasizes contextual performance, organizational citizenship, and pro- versus anti-social behavior, all of which concern getting along at work. Tett, Guterman, Bleier, and Murphy (2000) describe a taxonomy that synthesizes 12 models of managerial performance. They identify 53 performance dimensions, the content of which suggests the well-known factors of structure and consideration (Bass, 1990; Fiedler, 1967; Fleishman, 1953). Managers who provide structure enable groups to get ahead; managers who are considerate facilitate the ability of groups to get along.

Finally, Bartram (2005) analyzed the structure of the universe of competencies, defined as “sets of behaviors that are instrumental to the delivery of desired results” (Bartram, Robertson, & Callinan, 2002, p. 7). He begins with two metaconcepts that correspond to “getting along” and “getting ahead.” He expands the metaconcepts to include eight broad competency factors—“the Great Eight.” Competencies that promote getting along include Supporting and Cooperating, Interacting and Presenting, Organizing and Executing, and Adapting and Coping; competencies that promote getting ahead include Leading and Deciding, Analyzing and

Interpreting, Creating and Conceptualizing, and Enterprising and Performing. Bartram's competencies overlap with the generalized work activities that Jeanneret, Borman, Kubisiak, and Hanson (1999) proposed as a comprehensive taxonomy of work behaviors required in the US economy.

Job Performance Domains, Their Competencies, and Their Measurement

McClelland and his colleagues (e.g., Boyatzis, 1982) introduced the concept of *competency*, which they defined as performance capabilities that distinguish effective from ineffective personnel. McClelland defined competencies empirically in terms of the requirements of particular job in particular contexts. This rigorous approach is rare in a field characterized by ad hoc competency models.

As noted above, we believe that every existing competency model can be organized in terms of a "domain model" first proposed by Warrenfeltz (1995). The model contains four domains: (a) intrapersonal skills; (b) interpersonal skills; (c) technical skills; and (d) leadership skills. Hogan and Warrenfeltz (2003) argue that these four domains form a natural, overlapping developmental sequence, with the later skills depending on the appropriate development of the earlier skills. These domains also form a hierarchy of trainability, in which the earlier skills are harder to train and the later skills are easier to train.

Intrapersonal Skills. Intrapersonal skills develop early in childhood and then have important consequences for career development in adulthood. Two components underlie the domain of intrapersonal skills. The first is core self-esteem (Erez & Judge, 2001; Judge & Bono, 2001), emotional security, or resiliency. People with core self-esteem are self-confident, even-tempered, and positive, they are not easily frustrated, and they bounce back quickly from reversals and disappointments. Persons who lack core self-esteem are self-critical, moody,

unhappy, easily frustrated, and need frequent reassurance. Core self-esteem is easily assessed using any well-validated personality measure of emotional stability from the FFM. More importantly, measures of core self-esteem predict a wide variety of career outcomes, including job satisfaction and performance evaluations (Judge & Bono, 2001).

The second component of intrapersonal skills concerns self-control. Self-controlled people follow rules and respect authority; they are conforming, socially appropriate, and easy to supervise. Persons with low self-control ignore rules and procedures; they are rebellious, refractory, and hard to supervise. Self-control is easily assessed using well-validated personality measures of FFM emotional stability and conscientiousness (Hogan & Hogan, 1989); self-control predicts a wide variety of career outcomes, including supervisors' ratings of satisfactoriness (Hogan & Holland, 2003).

Intrapersonal skill is the foundation on which careers develop. Persons with good intrapersonal skills project integrity; integrity is perhaps the most important characteristic in employability. It is prerequisite for getting along with others because persons who act with integrity gain the reputation of being responsible, dependable, and trustworthy.

Interpersonal Skills. Interpersonal skills concern building and sustaining relationships. People with good interpersonal skills seem socially adept, approachable, and rewarding to deal with. Hogan and Warrenfeltz (2003) describe interpersonal skills in terms of three components. The first is a disposition to put oneself in the place of another person, to anticipate how that person sees the world and what he/she expects during an interaction. The second component involves getting it right when one does this. The third component involves incorporating the information about the other person's expectations into one's subsequent behavior.

Interpersonal skill concerns building and maintaining relationships with a variety of people who might differ from oneself in terms of demographic and psychological characteristics. Interpersonal skill is easily measured using well-validated personality measures of FFM extraversion and agreeableness (Bartram, 2005). Good measures of interpersonal skill predict a wide range of occupational outcomes, including supervisory performance (cf. J. Hogan & Hogan, 2001; Riggio, 1989). These skills are prerequisite for getting along with others because they are the foundation for establishing and sustaining relationships.

Technical Skills. The domain of technical skills is included in every comprehensive model of performance; it differs from the preceding two domains in several ways. Technical skills can be taught, and they are the least dependent on being able to deal with other people. Technical skills involve comparing, copying, compiling, computing analyzing, coordinating, innovating, synthesizing, and so on (Peterson, Mumford, Borman, Jeanneret, & Fleishman, 1999). These skills can be assessed using work simulations, assessment center exercises, and content-valid tests; the best predictors of individual differences in technical skills are measures of cognitive ability (Hunter & Hunter, 1984; Ree & Earles, 1992).

Interest in training and acquiring new technical knowledge is also part of this domain. The tendency to value education can be assessed using well-validated measures of the FFM dimension of openness to experience. These measures predict technical outcomes such as supervisor ratings of judgment, market savvy, training progress, and job knowledge (Hogan & Holland, 2003). Technical skills are prerequisite for getting ahead because persons who seem knowledgeable and competent are a resource for the performance of their group.

Leadership Skills. Leadership skills are relevant for performance in virtually any job. Leadership skills, which concern building and maintaining effective teams, can be understood in terms of five components—which depend on intrapersonal, interpersonal, and technical skills.

The first component is recruiting talented people to a team. The second component involves retaining that talent once it is recruited. The third component of leadership skills concerns motivating a team; other things being equal, a motivated team will outperform a talented but less motivated group. Recruiting, retaining, and motivating team members depend on building positive relationships with each team member, a capability that builds on the interpersonal skills described above. The fourth component concerns developing and promoting a vision for the team. The vision legitimizes the team enterprise. Technical competence is needed to develop the vision, and interpersonal skills are needed to sell it. The final component of leadership skill concerns being persistent and hard to discourage. Persistence probably depends on core self-esteem and conscientiousness, although there is little research on the topic.

Leadership skill can be assessed using any number of well-validated procedures, although the most effective assessment uses a combination of methods. Historically, assessment centers have been used extensively. Recent meta-analyses indicate that measures of cognitive ability (Judge, Colbert, & Ilies, 2004) and personality substantially predict both leadership emergence and effectiveness (Judge, Bono, Ilies, & Gerhardt, 2002). Leadership skills are prerequisite for getting ahead because leadership by definition involves the pursuit of status.

Summary

We use the domain model to organize the content of performance, and its measurement in VG research. The metaconcepts (getting along and getting ahead) and four competency domains are inclusive and exhaustive; they also resemble Hough and colleagues' notion of taxons (Hough

& Ones, 2001; Hough & Schneider 1996). Both approaches are flexible and inclusive, and in VG research, flexibility and inclusiveness are essential because every organization views its jobs as unique.

Converging Validity Generalization Evidence

The *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999) and the *Principles* use the term VG to refer to meta-analysis, transportability of validity, and synthetic/job component validity. These three techniques are used to evaluate the validity of selection procedures in particular cases. Although best practice dictates using multiple sources of validity evidence to evaluate a test (Binning & Barrett, 1989; Messick, 1995), we know little about using multiple VG models for this purpose. Projects by Hoffman, McPhail, and their colleagues (e.g., Hoffman, Holden, & Gale, 2000, Hoffman & McPhail, 1998) and Johnson, Carter, and their colleagues (e.g., Johnson, Carter, Davison, & Oliver, 2001), which used cognitive measures for VG across several job families in a single organization, are notable exceptions. Combining validity information from various sources requires technical expertise and professional judgment (Hoffman et al., 2000; Landy, 1986), but doing so should yield more defensible selection procedures than using a single strategy as endorsed by the *Uniform Guidelines on Employee Selection Procedures (UGESP: Equal Opportunity Employment Commission, 1978)*.

We generalize validity evidence from the HPI using the domain model (Table 1) and the Hogan Archive, which contains over 200 validity studies spanning 25 years. Meta-analysis and synthetic/job component validity research depend on using archival validity data that can be aggregated. VG research is appropriate when resources preclude local validation, when an

organization needs to hire immediately for newly created jobs, or when many jobs contain only a single incumbent. The research proceeds through five steps:

1. Analyze the personality-based job requirements for the marker job, and use the results to guide the subsequent study.
2. Aggregate meta-analytic validity evidence for the marker job from the published literature and the Hogan Archive.
3. Conduct transportability of validity research using criterion-related validity data from the “original study” in the Hogan Archive to the “new location” as required by the *UGESP*.
4. Conduct synthetic validity/job component validity research using competency ratings, examples of which are illustrated in Table 1. Using the Hogan Archive, we meta-analyze correlations between personality predictors and competency ratings, then use the results to define and evaluate a test battery.
5. We examine the validity evidence from the three VG sources, compose a battery for the new job, and simulate the consequences of various cutoff scores using archival data.

In the following sections, we describe go through these steps using managerial jobs as an example of how to conduct personality-based VG studies.

Step 1. Analyze Personality-Based Requirements for the Marker Job

The *UGESP* emphasizes that all validity studies should be based on a job analysis or review. The guidelines require documenting: (a) the required work behaviors and/or outcomes; (b) the criticality of the work behaviors or outcomes; and (c), if applicable, the evidence and rationale for grouping two or more jobs. Tippins and McPhail (2002) describe these requirements in detail. Generalizing validity requires information about the job for which validity evidence is available and the job to which one wants to generalize. The goal is to

evaluate the similarity of the two jobs. We use standardized job analysis methods to determine similarity; this allows us to estimate the defensibility and the possible error associated with a VG analysis. We use a mixed-method, multiple step job analysis as follows:

1. Review existing information about the marker job.
2. Code the important dimensions from the job descriptions.
3. Match the job to existing databases, including the O*NET (Peterson et al., 1999) descriptions and DoL job family typology.
4. Ask subject matter experts (SMEs) to complete our worker-oriented, personality-based job analysis, the Job Evaluation Tool (JET: Hogan Assessment Systems, 2000).
5. Conduct job observations, interviews, and critical incident focus groups with incumbents.

Typically, organizations have position descriptions for the marker job that contain useful information. We also use O*NET codes to determine the similarity of jobs and the JET is rationally linked to O*NET classifications. D'Egidio (2001) also provides empirical links between the HPI and O*NET ratings.

Most people assume that job analysis is a well-defined methodology, Cornelius (1988, p. 48) concluded that it is not; based on decades of research, he suggested that the best way to identify “necessary job traits” is to review “mental, social, and physical” traits using the Position Analysis Questionnaire (PAQ: McCormick, Jeanneret, & Mecham, 1972). Dunnette and Borman (1979) and Guion (1998) refer to the PAQ as a milestone in worker-oriented job analysis and the PAQ is good way to understand the cognitive requirements of a job. But the PAQ provides little information about a job’s non-cognitive demands. Other than Raymark, Schmit, and Guion’s (1997) *Personality-Related Position Requirements Form*, the non-cognitive domain is not emphasized in any well-established, standardized job analysis method. We

developed the *Performance Improvement Characteristics* (PIC) checklist to identify personality-based job requirements; it is part of the job analysis questionnaire we refer to as the *Job Evaluation Tool* (JET).

We use the JET to determine the critical requirements of a job in terms of (a) the personal capabilities that would improve performance, and (b) the critical competencies required for effective performance. The JET contains five sections covering various taxonomies of job performance; sample items from the JET are included in the Appendix. We use the *Performance Improvement Characteristics* and the *Competency Evaluation Tool* (CET) sections of the JET for personality-based VG research. They evaluate the competencies and personality requirements of the domain model (Table 1).

Quantitative Job Analysis. The PIC identifies (a) the personal characteristics (see measurement constructs in Table 1) needed for a job, and (b) the degree to which performance on the job is enhanced by them (Hogan & Rybicki, 1998). SMEs rate jobs on scales ranging from 0 (*Does Not Improve Performance*) to 3 (*Substantially Improves Performance*). Harvey (1991) recommends expressing job analysis methods and selection constructs in common terms; thus, the 48 PIC items form seven scales that align with the seven HPI scales (see Table 2). Internal consistency reliability estimates for the PIC scales range between .76 (Adjustment) and .87 (Interpersonal Sensitivity); the average is .83. Test-retest reliability, estimated over a one-month interval, ranges between .60 (Learning Approach) and .84 (Inquisitive); average test-retest reliability is .71.

We obtain PIC scale scores by (a) summing the item ratings on each of the seven scales, (b) averaging the scores for each scale across raters, and (c) converting the average scale scores to percentiles using a national database. The raw scores for each scale and their standard errors

of measurement are used to compare PIC profiles across different jobs. Previous research indicates that the PIC differentiates among jobs, and scores on PIC scales correlate with the HPI scales that predict job performance (Foster & Anderson, 2005; Rybicki, 1997; Tett, Holland, Hogan, & Burnett, 2002). The results of these studies support the validity of the PIC results for job analyses across different jobs and organizations (cf. Harvey & Wilson, 2000).

SMEs use the CET portion of the JET to rate the criticality of 40 competencies for successful job performance (see Table 1). SMEs evaluate the importance of each competency for the target job using a five-point scale, where “0” is defined as *Not associated with job performance* and “4” is defined as *Critical to job performance*. The 40 CET competencies map onto the four performance domains presented in Table 1, as well as onto other competency and performance models described previously in this chapter.

Internal consistency reliability estimates for domain scores, using ratings across 951 archived jobs for the 40 CET items, are .76 for Intrapersonal, .72 for Interpersonal, .77 for Technical, and .89 for Leadership; the average is .78. Average SME inter-rater reliability within a job was estimated using intraclass correlations (Shrout & Fleiss, 1979). Across 35 jobs, average inter-rater reliability for is .44 for a single rater, and .82 for a set of seven raters (Hogan Assessment Systems, 2006).

Job Analysis Results for a Marker Managerial Job: Case Study. A multinational telecommunications company wanted to improve its selection procedures for entry-level management jobs because the jobs are the starting point for management succession. The organization thought its current selection system could be enhanced by using valid personality assessments. The organization provided a description of a job corresponding to 1.1-1021.00 “General and Operations Managers” O*NET code. The O*NET description is:

Plan, direct, or coordinate the operations of companies or public and private sector organizations. Duties and responsibilities include formulating policies, managing daily operations, and planning the use of materials and human resources, but are too diverse and general in nature to be classified in any one functional area of management or administration, such as personnel, purchasing, or administrative services. Includes owners and managers who head small business establishments whose duties are primarily managerial.

SMEs (N = 225) described the job using the PIC and CET sections of the JET. SMEs had an average of 6.3 years experience supervising the job they rated, and 3.2 years experience as incumbents in these jobs. SMEs rated a representative sample of management-level jobs. The PIC results defined the personal characteristics associated with effective job performance. As seen in Figure 1, these included leadership (Ambition), resilience (Adjustment), conscientiousness (Prudence), and staying up-to-date (Learning Approach). Enjoying social interaction (Sociability) was less important for the job. Note that this profile combines elements of getting along and getting ahead, which are important for success as a manager.

SMEs also evaluated job competencies using the CET. Descriptive statistics for these ratings appear in Table 3. Average inter-rater reliability across the 40 CET competencies was .89. Thirteen competencies had mean ratings greater than “3”, which is interpreted as *Important to job performance*. They suggest that effective managers must act with integrity, make timely and well-reasoned decisions, be technically skilled, and provide vision—themes that appear in the job descriptions for management-level employees and are consistent with implicit leadership theory (Lord, DeVader, & Alliger, 1986).

Step 2. Meta-Analytic Evidence from Validity Studies for the Marker Job

The *Principles* (SIOP, 2003) recognize meta-analysis as a method “that can be used to determine the degree to which predictor-criterion relationships are specific to the situations in which the validity data have been gathered or are generalizable to other situations, as well as to

determine the sources of cross-situation variability (Aguinis & Pierce, 1998)” (p. 28). Pearson (1904; as cited in Rosenthal & DiMatteo, 2001) reported meta-analytic results evaluating the efficacy of vaccinations over 100 years ago. However, the method was only used to evaluate selection test validity in the late 1970s, and it was not the first method to be used (cf. Lawshe, 1952). Of the three VG methods, meta-analysis provides the most generalizable results, but it relies exclusively on criterion-related validity studies. Transportability and synthetic/job component validity research is less generalizable, but can use either content or criterion-related research as source data.

Meta-analysis averages findings from multiple studies of the same relationship to provide a best estimate of ρ (i.e., the correlation in the population) by controlling for error due to sampling, measurement range restriction, and moderators (Smith & Glass, 1977). In addition, there are standardized criteria for deciding what studies to include, what variables to code, effect size comparisons, and moderator identification. Ideally, a meta-analysis includes all relevant studies; however, this is often impossible because studies with insignificant results are less likely to be published. Rosenthal (1979) notes that this is a problem for meta-analytic research based on few studies, small sample sizes, and an atheoretical base.

According to the *Principles*, “reliance on meta-analysis results is more straightforward when they are organized around a construct or set of constructs” (p. 30). Schmidt and Hunter (1977) used a construct orientation in their well-known meta-analysis of cognitive ability measures. Hogan and Holland (2003) did the same in a meta-analysis of the validity of personality predictors (see Table 1). A construct driven approach has two advantages. First, theory drives professional judgment, which is unavoidable when compiling data from multiple studies. Second, a theory-driven approach provides a framework for interpreting the results.

The remainder of this section describes a meta-analysis of personality measures and managerial jobs.

Meta-Analysis of Personality Measures for a Managerial Job: Case Study. We proceeded in two phases. First, we identified a set of studies whose predictors and criteria are relevant for the job family. Several meta-analyses have focused on managerial jobs. Barrick and Mount (1991) found that FFM factors Conscientiousness ($\rho = .22$) and Extraversion ($\rho = .18$) predicted all managerial performance criteria (i.e., job proficiency, training proficiency, and personnel data). Barrick and Mount (1993) reported once again that Conscientiousness ($r = .25$) and Extraversion ($r = .14$) were significantly related to job performance. More recently, Hurtz and Donovan (2000) found that Conscientiousness ($\rho = .17$), Extraversion ($\rho = .12$), and Emotional Stability ($\rho = .12$) reliably predict managerial performance. Focusing on leadership, Judge, Bono, Ilies, and Gerhardt (2002) found that Emotional Stability ($\rho = .24$), Extraversion ($\rho = .31$), Openness to Experience ($\rho = .24$), and Conscientiousness ($\rho = .28$) predict performance. The most generalizable measure across samples and criteria was Extraversion (which includes Ambition). Finally, Larson, Rottinghaus, and Borgen (2002) and Barrick, Mount, and Gupta (2003) examined the relationship between the FFM and Holland's RIASEC occupational types. Extraversion (which includes Ambition) predicted ($\rho = .41$) interest in persuading and leading others to reach organizational or economic goals. Together, these analyses suggest that Emotional Stability, Extraversion, Conscientiousness, and Openness to Experience predict performance in the managerial-level job family.

Second, we searched the Hogan Archive for studies that include supervisor as well as entry-level, middle, and executive management jobs (Foster & Hogan, 2005), and identified 35 such studies. Each study reported correlations between the HPI scales and job performance

criteria. We aggregated the correlations for each scale across studies, using meta-analysis (Hunter & Schmidt, 1990) to estimate the true relationship between the predictor variables and job performance.

All studies used zero-order product-moment correlations. We corrected for sampling error, unreliability in the measures, and range restriction. We estimated the reliability of the personality measures using within-study coefficient alpha [$M = .78$; range = .71 (Prudence) to .84 (Adjustment)], rather than using the values reported in the HPI manual. Following Barrick and Mount (1991) and Tett et al. (1991), we used the .508 reliability coefficient proposed by Rothstein (1990) to estimate the reliability of supervisory ratings of job performance. We also computed a range restriction index for HPI scales. Following Hunter and Schmidt (1990), we divided each HPI scale's within-study standard deviation by the standard deviation reported in the test manual. This yields an index of range restriction for each HPI scale for each study. We used mean replacement within job family to estimate range restriction correction factors for each scale, when a within study standard deviation was unavailable.

We averaged correlations within studies so that each sample contributed only one estimate per scale. Note that this procedure uses negative and positive correlations rather than absolute values for averaging correlations, which differs from the analyses reported by Tett et al. (1991, p. 712). We did not correct correlations to estimate validity at the construct level. Although some (e.g., Mount & Barrick, 1995; Ones, Viswesvaran & Schmidt, 1993) argue this artifact can be corrected, we believe it is premature to estimate the validity of the perfect construct when there is no agreement about the definition of the perfect construct.

Table 4 contains the correlations for managerial performance. Consistent with previous research, HPI Adjustment and Ambition (FFM Emotional Stability and Extraversion [in part])

are the best predictors of performance. Again consistent with earlier results, Prudence and Interpersonal Sensitivity (FFM Conscientiousness and Agreeableness) predict overall performance, but less well. Finally, Sociability, Inquisitiveness, and Learning Approach (FFM Extraversion [in part] and Openness) had positive but small correlations with job performance, suggesting that they might be important in some specific cases.

Every meta-analytic study reviewed here (i.e., Barrick & Mount, 1991; 1993; Barrick et al., 2003; Foster & Hogan, 2005; Hurtz & Donovan, 2000; Judge et al., 2003; Larson et al., 2002) supports the use of personality measures for selection into management jobs. However, each study provides a slightly different estimate of the predictor-criterion relationships because they use different measures. Our meta-analysis (Foster & Hogan, 2005) used a single measure of the FFM (the HPI) and is consistent with the results from earlier published studies; consequently, we conclude that the HPI Adjustment ($r = .23$), Ambition ($r = .31$), Interpersonal Sensitivity ($r = .15$), Prudence ($r = .13$) and Learning Approach ($r = .10$) scales are the best predictors of managerial job performance. We used Nunnally's (1978) equation (cf. Johnson, Carter, and Tippins, 2001) to estimate the relationship between these scales and overall performance:

$$r_{y_c, x_c} = \frac{\overline{r_{y_i, x_i}}}{\sqrt{\overline{r_{yy}}} \sqrt{\overline{r_{xx}}}}$$

Based on this estimate, the overall validity of the test battery is $r = .41$.

Step 3. Transportability of Validity from Personality Predictors

The next step in the VG process is to transport validity evidence from one job to another similar job. The primary reference for the transportability of validity is the *UGESP*. Hoffman, McPhail and colleagues (Hoffman & McPhail, 1998; Tippins, McPhail, Hoffman, & Gibson,

1999) discuss the technical requirements for test transportability in situations that preclude local validation and use the *UGESP* as a reference. Johnson and Jolly (2000) provide an empirical demonstration of the method and note the lack of guidance for its appropriate use.

The *Principles* considers transportability of validity as one of three VG strategies for justifying a selection procedure. Transportability involves using a selection procedure in a new situation based on evidence established in another situation. This assumes that the original validation study was technically sound and that the jobs are “closely related” (Bernard v. Gulf Oil Corp., 1989). Situations where transportability might apply include choosing a selection procedure for the same job across multiple locations and different companies or for different jobs with similar requirements. It might also be useful for different job titles within a single job family.

The *UGESP*, the *Standards*, and the *Principles* all recognize transportability of selection procedures (cf. Tippins, 2003). Although employment discrimination experts distinguish between these three documents, we focus on their common themes. For example, all three require that the original research be technically adequate. *UGESP* emphasizes the need for evidence regarding fairness as well as validity and job similarity as criteria for transportability. Personality-based selection procedures typically yield no adverse impact, which meet requirements set by *UGESP* and precedents in many courts (Lindemann & Grossman, 1996). However, fairness is generally considered a social rather than a psychometric issue.

The *Standards* emphasize the need for good cumulative research (e.g., meta-analysis) and discourage relying on single local validation studies unless they are “exceptionally sound.” Interestingly, the original design for transportability of a selection procedure relies on a single validation study. The *Principles* emphasize the importance of job similarity in the original and

the new job. Evidence of similarity comes from job requirements, job context, and job applicants. For personality-based selection systems, demonstrating job similarity has been a problem because few job analysis methods were available. Notable exceptions are Raymark et al. (1997) and the JET methodology.

Job similarity is the key to transporting test validity. The success of the strategy rests on: (1) establishing validity in a marker job; and then (2) demonstrating that the marker job is comparable to the job to which the validity evidence is to be transported. We estimate similarity using converging evidence and professional judgment.

Transportability of Validity for a Managerial Job: Case Study. We conduct transportability of validity research in two steps. First, we analyze the “new location” target managerial job in qualitative and quantitative terms. Second, we identify an “original location” managerial job for which there is validation study. In the present case, we identified the job in the Hogan Archive based on identical O*NET codes, similar PIC profiles (see Figure 1), overlapping competencies from the CET, and similar applicant pools. The O*NET typology provided a standard external metric for rating job similarity.

We evaluate the similarity of the new job to the job in the Hogan archive using PIC profiles and CET importance ratings. We determine the similarity of PIC profiles by calculating the standard error of measure (SEM) for each of the seven PIC scales based on the coefficient alpha for each using the following formula:

$$SEM = s_t \sqrt{1 - r_{tt}}$$

Where s_t is the standard deviation of the PIC scale and r_{tt} is the reliability of the scale.

We construct a 95% confidence interval for each scale by adding and subtracting 1.96 SEMs to and from each raw score scale mean. We compare each PIC scale mean for the target

job with the confidence interval for the same scale for the marker job. The jobs are sufficiently similar for transportability of validity if the seven scale means for the target job fall within the 95% confidence intervals. Figure 1 presents PIC profiles with each scale mean, defined as a percentile calculated from the raw score mean divided by total possible raw score for the scale; we calculate the confidence interval for each scale in the same way.

Overlap analyses of CET ratings provide further evidence of job similarity. In this example, thirteen of the fourteen competencies rated as important for the target job were also rated as important for the marker job. We also computed the overlap statistic referenced by Fleishman & Quaintance (1994):

$$N_s / N_s + N_d$$

Where N_s (i.e., $n = 41$) is the number of competencies evaluated in the same way for both the target and archive jobs (i.e., either both important or both not important), and N_d (i.e., $n = 15$) is the number evaluated differently. This overlap statistic was .73. Moreover, the correlation between ratings for the two jobs on the 20 most important competencies was .96. Finally, mean differences across jobs for competencies rated as important averaged .21, and ranged from .05 to .3 on the 0 – 4 scale; this indicates substantial agreement among the SMEs regarding the similarity of the two jobs.

We judged similarity of applicant pools across the target and marker jobs using the demographics of the SME samples in the job analysis, the incumbent samples in the criterion studies, and the subsequent applicant pool for the positions. Specifically, the samples for each job were primarily white (i.e., 64% and 74%) and male (i.e., 69% and 71%). We also compare mean HPI scale scores for the two incumbent samples; mean differences between the two groups for each scale were within one standard error of measure.

Based on this converging evidence on job similarity, the similarity of the potential applicant pools, and the existence of an adverse impact study, we determined that Study #324 from the Hogan Archive was appropriate for transporting validation evidence. Table 5 presents correlations between the HPI scales and most important competencies for Study #324. The average corrected correlations between the HPI and the nine competencies for this managerial sample are as follows: Adjustment (.34); Ambition (.20), Interpersonal Sensitivity (.14), and Prudence (.16). The similarity of the jobs indicates that the test battery can be transported and used for selection into the new job. Nunnally's (1978) equation suggests that the validity of the test battery for predicting overall job performance from the transportability of validity results is .30.

Step 4. Synthetic Validity/Job Component Validity for Personality Measures

The most specific validity generalizability evidence comes from synthetic validity/job component validity research. Mossholder and Arvey (1984) noted that, whereas meta-analysis relies on global evaluations of job similarity, synthetic validity requires a detailed examination of the work. The strategy is criterion driven and the selection procedure resembles using canonical correlations—that is, finding the best set of predictors to maximize the best representation of the criterion space. The three VG methods fill the gaps inherent in each.

The *UGESP* is vague about technical requirements and documentation for synthetic/job component validity, but the *Principles* explicitly include this strategy. Lawshe (1952) introduced synthetic validity over 50 years ago; however, it was largely ignored because people believed that test validity is specific to situations. The interpretive review and demonstration by Mossholder and Arvey (1984) is a rare exception. Mossholder and Arvey define synthetic validity as “the logical process of inferring test-battery validity from predetermined validities of

the tests for basic work components” (p. 323). If we know the key components of a job, we can review studies predicting those components. The valid predictors of the key job components can then be “synthesized” into a test battery for the new job (Balma, 1959; Lawshe, 1952). Primoff (1959), Guion (1965), and McCormick, DeNisi, and Shaw (1979) also examine synthetic validity. Hoffman, Holden, and Gale (2000), Jeanneret and Strong (2003), Johnson, Carter, Davison, and Oliver (2001), and McCloy (1994, 2001) have published synthetic validity research, and Scherbaum (2005) reviews of the field. Brannick and Levine (2002) point out that synthetic validity approaches allow us to build up validity evidence from small samples with common job components.

Synthetic validation involves: (1) identifying the important components of a job; (2) reviewing prior research on the prediction of each component; and (3) aggregating correlations across multiple studies for each component of the job to form a test battery (Scherbaum, 2005). Mossholder and Arvey (1984) summarize these requirements as follows: “When test battery validity is inferred from evidence showing that tests measure broad characteristics necessary for job performance, the process resembles a construct validation strategy. When scores are correlated with component performance measures, the process involves criterion-related validation. The nature of the tests used in the process (e.g., work sample vs. aptitude) may determine in part the appropriate validation strategy” (p. 323).

One version of synthetic validation is job component validity (JCV: McCormick, DeNisi, & Shaw, 1979). JCV has been used primarily to study the cognitive demands of jobs by correlating job dimensions using PAQ data (Jeanneret, 1992; Hoffman, et al., this volume). Jeanneret describes JCV as falling “within the rubric of construct validity” and reviews evidence from studies using cognitive ability tests. Hoffman and McPhail (1998) examined the accuracy

of JCV for predicting the observed validity of cognitive tests in clerical jobs. Few similar analyses are available for personality predictors, although Mecham (1985) and D'Egidio (2001) provide notable exceptions.

Synthetic/Job Component Validity for a Managerial Job: Case Study.

As pointed out by Guion (1965), Hoffman and McPhail (1998), McCormick (1959), and Mossholder and Arvey (1984), the first step in synthetic validation is a job analysis to determine the important components of the job. Our example is the managerial position analyzed above, and we use the CET results from the job analysis for the synthetic part of the study. Table 3 contains descriptive statistics for the 13 job dimensions the SMEs rated as important or critical for the target job.

Next we identify validity evidence for the important job components. Foster and Hogan (2005) have mapped each of the criteria from over 200 validity studies in the Hogan Archive onto the CET dimensions. We then conduct a meta-analysis for each scale by competency relationship, using all of the studies that included each competency. This meta-analysis provides stable estimates of the relationships between the seven HPI scales and the 13 critical competencies as rated by SMEs. This information is presented in Table 6. The correlations, averaged across the seven HPI scales and the 13 competencies, show that Adjustment (.14), Ambition (.20), Interpersonal Sensitivity (.09), and Prudence (.09) are stable predictors of the important competencies of the manager job.

Finally, we combine the validities across predictors and job components into a single coefficient representing the link between the predictor battery and total job performance. There are several methods for doing this and they are reviewed by Scherbaum (2005). Peterson, Wise, Arabian, & Hoffman (2001) specifically discuss various weighting options for predictor

batteries. Although these authors find little difference in the outcomes of the various methods, there are differences in data requirements (e.g., need for PAQ data). The data in the target study (i.e., CET) dictated that we use the weighting procedure recommended by Johnson, Carter, and Tippins (2001). Using the equation provided by Nunnally (1978), we calculate that the overall synthetic correlation a composite of the predictors and a composite of the critical components for the marker job is .26. This evidence supports the use of the HPI Adjustment, Ambition, Interpersonal Sensitivity, and Prudence scales to predict performance in the target job.

Step 5. Examine Validity Evidence and Simulate Outcomes from the Selection Procedure

The final step involves merging the evidence from each VG method into a single profile for the target job. This step is critical, but with the possible exception of Hoffman et al's (2000) work, the process is poorly described in the literature. Table 7 presents the results from the three VG strategies for the managerial job used in this chapter, summarized by predictor and validity source. The meta-analysis and the transportability of validity results support the use of five HPI scales for manager selection: Ambition, Adjustment, Prudence, Interpersonal Sensitivity, and Learning Approach. Evidence from the synthetic validity study only justifies using the Adjustment and Ambition scales. On balance, however, the data in Table 7 support using the Ambition, Adjustment, Prudence, Interpersonal Sensitivity, and Learning Approach scales as a test battery. This conclusion is also consistent with the overall corrected validity estimates (Nunnally, 1978) for the test battery. The mean of these corrected validity estimates is $r = .36$. Given the number of studies ($n > 200$) used to calculate this correlation, it is the best estimate of the overall validity of the HPI for predicting managerial performance.

Setting cutoff scores for a selection system depends on three criteria: (a) the validity of each scale, (b) the client's desired pass rate based on business necessity, and (c) the results of an

adverse impact study using the HPI normative data set ($n = 161,000$). Table 7 suggests that the Adjustment and Ambition scales are the best predictors of managerial performance. This conclusion translates into higher cutoff scores on these scales. The combined validity evidence for the remaining three scales suggests similar, but lower, cutoff scores for them. Based on our experience using these scales with similar applicant pools and the desired pass rate, we propose initial cutoff scores. We then conduct a simulation, using the HPI normative sample, to evaluate the effects of the cutoff scores on a potential applicant pool. We adjust the cutoff scores to yield the desired pass rate. At this point, we conduct adverse impact analyses with these cut scores. If adverse impact for any group is detected, based on the *UGESP* four-fifths rule, we modify the cutoff scores until it is minimized. Once the data become available, we compare our pass rate estimates with those from the actual applicant pool. These profiles have yet to result in adverse impact with an actual applicant pool.

Discussion

In this chapter, we described how three VG approaches may be used to evaluate the validity of personality measures within a competency domain framework. The purpose of the exercise was not to determine which of these methods works best, but to illustrate how they compliment each other. We believe that these three validity estimates provide the best solution for generalizing the validity of a personality-based selection procedure to a new job. This solution simply extends the unitarian view of validity (Binning & Barrett, 1989; Landy, 1986; Messick, 1995) to selection VG. We now comment briefly on each approach, and then consider outstanding future issues.

Conditions for Meta-Analysis, Transportability, and Synthetic/Job Component Validity

The availability of data usually dictates the most appropriate VG method. We agree with Hoffman and McPhail (1998) that VG should not rely entirely on meta-analysis, and we say this for two reasons. First, meta-analysis does not require job analysis—which is critical for real world test selection. Second, meta-analysis results may fail to detect less powerful but nonetheless important predictors of job performance, as seen in this chapter. Nevertheless, many researchers (e.g., Schmidt & Hunter, 1977; Peterson et al, 2001) believe that meta-analytic evidence is sufficient for VG as long as the predictor-criterion relationships are consistent across all components of job performance (i.e., no moderators identified).

Hoffman and McPhail (1998) note that transportability of validity is only practical in cases where “the test publisher has shown the forethought and taken the necessary steps to build-in this option while conducting their original validation research” (p. 988). The necessary steps to which they refer include archiving standardized job analysis information and empirical validation research on the original, marker, jobs. These two steps are essential for estimating the transportability of validity, but few test publishers maintain the requisite archives.

Synthetic/job component validation research requires data similar to the transportability of validity approach (i.e., standardized job analysis and archived validity coefficients), as well as validity data archived as job component criteria (Peterson et al., 2001). These requirements are both the strength of and the challenge for the synthetic validity method. If data are available to conduct a synthetic validation study, the approach is useful for virtually any job. We developed the CET job analysis tool to facilitate synthetic validity studies and the results from the CET are stored in a searchable database. Furthermore, each synthetic/job component validity study uses meta-analytic estimates of job component validities from which a personality predictor battery

can be compiled. Bartram (2005) illustrates this procedure using the Inventory of Management Competencies (SHL Group, 1993) and a database of 29 validation studies based on the Occupational Personality Questionnaire family.

Regardless of the validity generalization method, we found good validity for personality measures. This is consistent with current meta-analytic research (e.g., Hogan & Holland, 2003; Judge et al., 2002), and challenges the view (cf. Murphy & Dzieweczynski, 2005) that personality measures lack practical validity. The results of the three validity generalization studies reported here are similar in magnitude to results using other types of assessments (Hoffman et al., 2000).

Legal Basis for Validity Generalization Approaches

Unlike VG for cognitive ability tests, we found no existing case law on VG for personality-based test batteries. This result may be due to the fact that personality measures used for selection don't result in adverse impact, or it may be due to the fact that personality-based VG is a new activity. Nonetheless, challenges will inevitably emerge. As soon as selection procedures to include both cognitive and personality measures, adverse impact is almost guaranteed, and the adequacy of VG methods will be scrutinized. Interestingly, Guion (1965) proposed using this is the kind of test battery over 40 years ago.

Despite the regulatory and research support for VG studies, some employers will want to augment them with local validation. This has at least two benefits. First, organizations can use their own data to support the validity of their selection procedures, thereby minimizing questions about the appropriateness of generalized research. The legal defensibility of a local validation study is appealing to employers. Second, local validation can provide evidence of utility that

otherwise would not be available through VG methods. Brannick (2001) provides a Bayesian approach for combining a local validation study with existing validity generalization evidence.

Future Research Needs for Personality-Based VG

Future personality-based VG studies could benefit from the example of JCV research using the PAQ and cognitive measures. This tradition illustrates the importance of a standardized job analysis database and an archive of validity studies across jobs, construct measures, and criteria. The next generation of VG should also expand the measurement domains covered by test batteries to map more comprehensively the domain of job performance. Our work with personality measures expands the measurement domain and extends the coverage to other important competencies.

Theory-based models based are essential for future progress. Once the FFM was recognized as an organizing framework, the use of personality measures for personnel selection progressed substantially. It is perhaps less well recognized that the FFM concerns acceptance (getting along) and status (getting ahead). This framework provides structure and organization to a variety of instruments that are relevant for personnel selection. These domains structure the class of predictors in a VG selection battery. As taxonomies such as the FFM, verbal and quantitative cognitive abilities, and Holland's six vocational interest factors become better defined and measurement improves, so will our VG capability.

The performance domain model in Table 1 is a practical guide for linking criterion and construct measures. The domain model needs to be expanded to include missing measures whose validity can be tested. Minimally, this includes motivational and ability measures whose relations with the competency criteria can be estimated across each domain. Bartram's (2005, p. 1187) mapping of hypothesized measures and competency domains is a good starting point for

developing a matrix of relations. In addition, the criterion competency list and its organization across the four performance domains needs additional technical and conceptual work.

For VG applications, we envision developing a matrix that provides a single system of empirically validated relations for predictors and their associated competencies. Such a system should operate at the level of job families; that is, we need to specify test batteries for the 22 DoL job families using meta-analysis. Once a new job is classified by job family, the test battery can be refined using competencies identified across the domain model. For each job competency, meta-analyzed predictor-criterion validity estimates should be calculated to obviate the need for expert judgment of relations. The initial work with this approach seems promising, although with an increasingly fine-grained analysis, “piece-meal” synthetic validity could become a daunting task (Steel, Huffcutt, & Kammeyer-Mueller, 2006, p. 31). Nevertheless, increasing the database for VG will advance our ability to deliver high-quality, comprehensive selection systems economically.

References

- Adler, A. (1939). *Social interest*. New York: Putnam.
- Aguinis, H., & Pierce, C. A. (1998). Testing moderator variable hypotheses meta-analytically. *Journal of Management*, *24*, 577–592.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Arthur, W., Jr., Day, E. A., McNelly, T. L., & Edens, P. S. (2003). A meta-analysis of the criterion-related validity of assessment center dimensions. *Personnel Psychology*, *56*, 125-154.
- Bakan, D. (1966). *The duality of human existence: Isolation and communion in Western man*. Boston: Beacon.
- Balma, M. J. (1959). The development of processes for indirect or synthetic validity. *Personnel Psychology*, *12*, 395-396
- 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. (1993). Autonomy as a moderator of the relationship between the Big Five personality dimensions and job performance. *Journal of Applied Psychology*, *78*, 111-118.
- Barrick, M. R., Mount, M. K., & Gupta, R. (2003). Meta-analysis of the relationship between the Five-Factor Model of personality and Holland's occupational types. *Personnel Psychology*, *56*, 45-74.
- 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.
- Bartram, D., Robertson, I. T., & Callinan, M. (2002). Introduction: A framework for examining organizational effectiveness. In I. T. Robertson, M. Callinan, & D. Bartram (Eds.), *Organizational effectiveness: The role of psychology* (pp. 1-10). Chichester, UK: Wiley.
- Bass, B. M. (1990). *Bass & Sogdill's Handbook of Leadership: Theory, research, and managerial applications*. New York: Free Press.

- Binning, J. F., & Barrett, G. V. (1989). Validity of personnel decisions: A conceptual analysis of the inferential and evidential bases. *Journal of Applied Psychology, 74*, 478-494.
- Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt, W. C. Borman, & Associates (Eds.), *Personnel selection in organizations* (pp. 71-98). San Francisco, CA: Jossey- Bass.
- Boyatzis, R. E. (1982). *The competent manager: A model for effective performance*. New York: Wiley.
- Brannick, M. T. (2001). Implications of empirical Bayes meta-analysis for test validation. *Journal of Applied Psychology, 86*, 468-480.
- Brannick, M. T., & Levine, E. L. (2002). *Doing a job analysis study*. In *job analysis: Methods, research, and applications for human resource management in the new millennium* (Ch. 9, pp. 265-294). Thousand Oaks, CA: Sage.
- Campbell, J. P., Gasser, M. B., & Oswald, F. L. (1996). The substantive nature of job performance variability. In K. R. Murphy (Ed.), *Individual Differences and Behavior in Organizations* (pp 258-299). San Francisco: Jossey-Bass.
- Campbell, J. P., McCloy, R. A., Oppler, S. H. & Sager, C. E. (1993). A theory of performance. In N. Schmitt, W. C. Borman & Associates (Eds.), *Personnel selection in organizations* (pp. 35-70). San Francisco: Jossey-Bass.
- Campbell, J. P., McHenry, J. J., & Wise, L. L. (1990). Modeling job performance in a population of jobs. *Personnel Psychology, 43*, 313-333.
- Cornelius, E. T. (1988). Practical findings from job analysis research. In S. Gael (Ed.), *The job analysis handbook for business, industry, and government*. New York: John Wiley & Sons.
- Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.
- D'Egidio, E. L. (2001). *Building a job component validity model using job analysis data from the Occupational Information Network*. Unpublished doctoral dissertation. University of Houston, Houston, TX.
- De Raad, B., & Perugini, M. (Eds.). (2002). *Big Five assessment*. Seattle, WA: Hogrefe & Huber.
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417-440.

- Dunnette, M. D., & Borman, W. (1979). Personnel and classification systems. In M. Rosenzweig & L. Porter (Eds.), *Annual review of psychology*, 30 (pp. 477-526).
- Emler, N. P. (1990). A social psychology of reputation. *European Review of Social Psychology*, 1, 173-193.
- Erez, A., & Judge, T. (2001). Relationship of core self-evaluations to goal setting, motivation, and performance. *Journal of Applied Psychology*, 86, 1270-1279.
- Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor, & Department of Justice. (1978). Uniform guidelines on employee selection procedures. *Federal Register*, 43, 38290-38315.
- Fiedler, F. E. (1967). *A theory of leadership effectiveness*. New York: McGraw-Hill.
- Fleishman, E. A. (1953). The measurement of leadership attitudes in industry. *Journal of Applied Psychology*, 37, 153-158.
- Foster, J. F., & Anderson, M. (2006, May). The validity of structured job analysis instruments. In J. Foster (Chair), Standardized job analysis tools: State of the science. Symposium conducted at the 21st Annual Conference of the Society for Industrial and Organizational Psychology, Dallas, TX.
- Foster, J. F., & Hogan, J. (2005). *Validity of the Hogan Personality Inventory for job family profiles*. Technical Report. Tulsa, OK: Hogan Assessment Systems.
- Gatewood, R. D., & Feild, H. S. (1994). *Human resource selection (3rd ed.)*. Orlando, FL: Dryden Press.
- Gaugler, B. B., Rosenthal, D. B., Thornton, G. C., & Bentson, C. (1987). Meta-analysis of assessment center validity. *Journal of Applied Psychology*, 72, 1-28.
- Ghiselli, E. E. (1966). *The validity of occupational aptitude tests*. New York: Wiley.
- Ghiselli, E. E., & Brown, C. H. (1955). *Personnel and industrial psychology (2nd ed.)*. New York: McGraw-Hill.
- Goffman, E. (1958). *The presentation of self in everyday life*. New York: Doubleday.
- Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*, 59, 1216-1229.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4, 26-42.

- Guion, R. M. (1965). Synthetic validity in a small company: A demonstration. *Personnel Psychology, 18*, 40-63.
- Guion, R. M. (1998). *Assessment, measurement, and prediction for personnel decisions*. Mahwah, NJ: Lawrence Erlbaum.
- Harvey, R. J. (1991). Job analysis. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2, 2nd ed., pp. 71-163). Palo Alto, CA: Consulting Psychologists Press.
- Harvey, R. J., & Wilson, M. A. (2000). Yes Virginia, there is an objective reality in job analysis. *Journal of Organizational Behavior, 21*, 829-854.
- Hoffman, C. C., Holden, L. M., & Gale, E. (2000). So many jobs, so little “n”: Applying expanded validation models to support generalization of cognitive ability. *Personnel Psychology, 53*, 955–991.
- Hoffman, C. C., & McPhail S. M. (1998). Exploring options for supporting test use in situations precluding local validation. *Personnel Psychology, 51*, 987–1003.
- Hogan Assessment Systems. (2000). *Job Evaluation Tool manual*. Tulsa, OK: Author.
- Hogan Assessment Systems. (2006). *Competency Evaluation Tool manual*. Tulsa, OK: Author.
- Hogan, J., & Hogan, R. (1989). How to measure employee reliability. *Journal of Applied Psychology, 74*, 273-279.
- 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., & Rybicki, S. (1998). *Performance Improvement Characteristics manual*. Tulsa, OK: Hogan Assessment Systems.
- Hogan, R. (1983). A socioanalytic theory of personality. In M. M. Page (Ed.), *1982 Nebraska symposium on motivation* (pp. 55-89). Lincoln: University of Nebraska Press.
- Hogan, R. (1986). *Manual for the Hogan Personality Inventory*. Minneapolis: National Computer Systems.
- Hogan, R. (1996). A socioanalytic perspective on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality* (pp. 163-179). New York: Guilford Press.
- Hogan, R., & Hogan, J. (1995). *Hogan Personality Inventory manual* (2nd ed.). Tulsa, OK: Hogan Assessment Systems.

- Hogan, R., & Hogan, J. (2001). Leadership and sociopolitical intelligence. In R. Riggio (Ed.), *Multiple intelligences and leadership* (pp. 65-89). San Francisco: Jossey-Bass.
- Hogan, R., & Shelton, D. (1998). A socioanalytic perspective on job performance. *Human Performance, 11*, 129-144.
- Hogan, R., & Warrenfeltz, W. (2003). Educating the modern manager. *Academy of Management Learning and Education, 2*, 74-84.
- Hough, L. M. (1992). The “Big-Five” personality variables—construct confusion: Description versus prediction. *Human Performance, 5*, 139-156.
- Hough, L. M., & Ones, D. S. (2001) The structure, measurement, validity, and use of personality variables in industrial, work, and organizational psychology. In N. R. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of work psychology* (pp. 233-277). London and New York: Sage.
- Hough, L. M., & Schneider, R. J. (1996). Personality traits, taxonomies, and applications in organizations. In K. R. Murphy (Ed.), *Individual differences and behavior in organizations* (pp. 31-88). San Francisco: Jossey-Bass.
- Hunt, S. T. (1996). Generic work behavior: An investigation into the dimensions of entry-level, hourly job performance. *Personnel Psychology, 49*, 51-83.
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96*, 72-98.
- Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis*. Newbury Park, CA: Sage.
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The big five revisited. *Journal of Applied Psychology, 85*, 869-879.
- Jeanneret, P. R. (1992). Applications of job component/synthetic validity to construct validity. *Human Performance, 5*, 81-96.
- Jeanneret, P. R., Borman, W. C., Kubisiak, U. C., & Hanson, M. A. (1999). Generalized work activities. In N. G. Peterson, M. D. Mumford, W. C. Borman, P. R. Jeanneret, & E. A. Fleishman (Eds.), *An occupational information system for the 21st century: The development of the O*NET* (pp. 105-125). Washington, DC: American Psychological Association.
- Jeanneret, P. R., & Strong, M. H. (2003). Linking O*Net job analysis information to job requirement predictors: An O*Net application. *Personnel Psychology, 56*, 465-492.

- 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.
- Johnson, J. W., Carter, G. W., Davison, H. K. & Oliver, D. H. (2001). A synthetic validity approach to testing differential prediction hypotheses. *Journal of Applied Psychology*, *86*, 774-780.
- Johnson, J. W., Carter, G. W., & Tippins, N.T. (2001, April). *A synthetic validity approach to the development of a selection system for multiple job families*. In J. Johnson, & G. Carter (Chairs), *Advances in the application of synthetic validity*. Symposium conducted at the 16th Annual Conference of the Society for Industrial and Organizational Psychology, San Diego, CA.
- Johnson, M. A., & Jolly, J. P. (2000). Extending test validation results from one plant location to another: Application of transportability evidence. *The Journal of Behavioral and Applied Management*, *1*, 127 – 136.
- Judge, T. A., & Bono, J. E. (2001). Relationship of core self-evaluations traits—self esteem, generalized self-efficacy, locus of control, and emotional stability—with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology* [Short Note], *86*, 80-92.
- 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., Colbert, A. E., & Ilies, R. (2004). Intelligence and leadership: A quantitative review and test of theoretical propositions. *Journal of Applied Psychology*, *89*, 542-552.
- Landy, F. J. (1986). Stamp collecting versus science: Validation as hypothesis testing. *American Psychologist*, *41*, 1183-1192.
- Larson, L. M., Rottinghaus, P. J., & Borgen, F. H. (2002). Meta-analyses of big six interests and big five personality factors. *Journal of Vocational Behavior*, *61*, 217-239.
- Lawshe, C. H. (1952). What can industrial psychology do for small business? (A symposium) 2. Employee selection. *Personnel Psychology*, *5*, 31-34.
- Lord, R. G., DeVader, C. L., & Alliger, G. (1986). A meta-analysis of the relation between personality traits and leader perceptions. *Journal of Applied Psychology*, *71*, 402-410.
- McCloy, R. A. (1994). Predicting job performance scores without performance data. In B. F. Green, & A. S. Mavor (Eds.), *Modeling cost and performance for military enlistment: Report of a workshop*. Washington, DC: National Academy Press.

- McCloy, R. A. (2001, April). *Predicting job performance scores in jobs lacking criterion data*. In J. Johnson, & G. Carter (Chairs), *Advances in the application of synthetic validity*. Symposium conducted at the 16th Annual Conference of the Society for Industrial and Organizational Psychology, San Diego, CA.
- McCormick, E. J. (1959). Application of job analysis to indirect validity. *Personnel Psychology, 12*, 402-413.
- McCormick, E. J., DeNisi, A. S., & Shaw, J. B. (1979). Use of the Position Analysis Questionnaire for establishing the job component validity of tests. *Journal of Applied Psychology, 64*, 51-56.
- McCormick, E. J., Jeanneret, P. R., & Mecham, R. C. (1972). A study of job characteristics and job dimensions based on the Position Analysis Questionnaire (PAQ). *Journal of Applied Psychology, 56*, 347-368.
- McCrae, R. R., & Costa, P. T., Jr. (1987). Validity of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*, 81-90.
- McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology, 86*, 730-740.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, 741-749.
- Mossholder, K. W., & Arvey, R. D. (1984). Synthetic validity: A conceptual and comparative review. *Journal of Applied Psychology, 69*, 322-333.
- Motowidlo, S. J., Borman, W. C., & Schmit, M. J. (1997). A theory of individual differences in task and contextual performance. *Human Performance, 10*, 71-83.
- Mount, M. K., & Barrick, M. R. (1995). The Big-Five personality dimensions: Implications for research and practice in human resource management. *Research in Personnel and Human Resource Management, 13*, 153-200.
- Mount, M. K., & Barrick, M. R. (2001). *Personal Characteristics Inventory manual*. Libertyville, IL: Wonderlic.
- Murphy, K. R., & Dzieweczynski, J. L. (2005). Why don't measures of broad dimensions of personality perform better as predictors of job performance? *Human Performance, 18*, 343-357.
- Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology, 66*, 574-583.

- Nunnally, J.C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
- Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validation: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology, 78*, 679-703.
- Ozer, D. J., & Benet-Martinez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology, 57*, 8.1-8.21.
- Peterson, N. G., Mumford, M. D., Borman, W. C., Jeanneret, P. R., & Fleishman, E. A. (1999). *An occupational information system for the 21st century: The development of the O*NET*. Washington, DC: American Psychological Association.
- Peterson, N. G., Wise, L. L., Arabian, J., & Hoffman, R. G. (2001). Synthetic validation and validity generalization: When empirical validation is not possible. In J. P. Campbell & D.J. Knapp (Eds.), *Exploring the limits in personnel selection and classification*. (pp. 411-451). Mahwah, NJ: Lawrence Erlbaum Associates.
- Primoff, E. S. (1959). Empirical validation of the J-coefficient. *Personnel Psychology, 12*, 413-418.
- Rank, O. (1945). *Will therapy and truth and reality*. New York: Knopf.
- Raymark, P. H., Schmit, M. J., & Guion, R. M. (1997). Identifying potentially useful personality constructs for employee selection. *Personnel Psychology, 50*, 723-736.
- Redfield, R. (1960). How society operates. In H. L. Shapiro (Ed.), *Man, culture, and society* (pp.345-368). New York: Oxford University Press.
- Ree, M. J., & Earles, J. A. (1992). Predicting training success: Not much more than g. *Personnel Psychology, 44*, 321-332.
- Riggio, R. E. (1989). *Social skills inventory manual*. Palo Alto, CA: Consulting Psychologists Press.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*, 638-641.
- Rosenthal, R., & DiMatteo, M. R. (2001). Meta analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52*, 59-82.
- Rothstein, H. R. (1990). Interrater reliability of job performance ratings: Growth to asymptote level with increasing opportunity to observe. *Journal of Applied Psychology, 75*, 322-327.

- Rotundo, M., & Sackett, P. R. (2002). The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: A policy capturing approach. *Journal of Applied Psychology, 87*, 66-80.
- Rybicki, S. (1997). *Validity of personality measures for entry-level jobs*. Tulsa, OK: Hogan Assessment Systems.
- Salgado, J. F. (1997). The five factor model of personality and job performance in the European community. *Journal of Applied Psychology, 82*, 30-43.
- Salgado, J. F. (1998). Big Five personality dimensions and job performance in Army and civil occupations: A European perspective. *Human Performance, 11*, 271-288.
- Scherbaum, C. A. (2005). Synthetic validity: Past, present, and future. *Personnel Psychology, 58*, 481-515.
- Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology, 62*, 529-540.
- Schmidt, F. L., & Rothstein, H. R. (1994). Applications of validity generalization methods of meta-analysis to biographical data scores in employees' selection. In G. S. Stokes, M. D. Mumford, & W. A. Owens (Eds.), *The biodata handbook: Theory, research, and applications* (pp. 237-260). Palo Alto, CA: Consulting Psychologists Press.
- SHL Group. (1993). *Inventory of Management Competencies: Manual and user's guide*. Thames Ditton, United Kingdom: Author.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing reliability. *Psychological Bulletin, 86*, 420-428.
- Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32*, 752-760.
- Society for Industrial and Organizational Psychology (2003). *Principles for the validation and use of personnel selection procedures* (4th ed.). Bowling Green, OH: Author.
- Steel, P. D. G., Huffcutt, A. I., & Kammeyer-Mueller, J. (2006). From the work one knows the worker: A systematic review of the challenges, solutions, and steps to creating synthetic validity. *International Journal of Selection and Assessment, 14*, 16-36.
- Tett, R. P., Guterman, H. A., Bleier, A., & Murphy, P. J. (2000). Development and content validation of a "hyperdimensional" taxonomy of managerial competence. *Human Performance, 13*, 205-251.

- Tett, R. P., Holland, B., Hogan, J., & Burnett, D. (2002, April). *Validity of trait-based job analysis using moderator correlations*. Paper presented at the 17th Annual Conference of the Society for Industrial and Organizational Psychology, Toronto, Ontario, Canada.
- 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.
- Tippins, N. T. (2003, October). *Transporting the validity of assessments*. Presentation at the annual meeting of International Assessment Congress, Atlanta, GA.
- Tippins, N. T., & McPhail, S. M. (2002, February). *New developments in transporting validity*. Presentation at PTC-MW education workshop, Washington, DC.
- Tippins, N.T., McPhail, S.M., Hoffman, C., & Gibson, W. (1999, April). *Transporting validity in the real world*. Continuing Education Workshop presented at the 14th Annual Conference of the Society of Industrial and Organizational Psychology, Atlanta, GA.
- Tupes, E. C., & Christal, R. E. (1961). *Recurrent personality factors based on trait ratings* (Tech. Rep. No. ASD-TR-61-97). Lackland Air Force Base, TX: Aeronautical Systems Division, Personnel Laboratory.
- Warrenfeltz, R. B. (1995, May). *An executive-level validation of the Borman and Brush taxonomy*. Paper presented at the 10th Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.
- Wiggins, J. S. (1996). *The Five-Factor Model of personality*. New York: Guilford.
- Wiggins, J. S., & Pincus, A. L. (1992). Personality structure and assessment. *Annual Review of Psychology, 43*, 473-504.
- Wiggins, J. S., & Trapnell, P. D. (1996). A dyadic-interactional perspective on the Five-Factor model. In J. S. Wiggins (Ed.), *The Five-Factor model of personality* (pp. 88-162). New York: Guildford.

Table 1

Domain Model of Job Performance, Example Competencies, and Personality Measures

Metaconcept	Domain	Example Competency	FFM Measurement	
Getting Ahead	Leadership	Achievement	Surgency/Extraversion	
		Building Teams		
		Business Acumen		
		Decision Making		
		Delegation		
		Employee Development		
		Initiative		
		Leadership		
		Managing Performance		
	Resource Management			
	Technical		Analysis	Openness to Experience
			Creating Knowledge	
			Decision Making	
			Political Awareness	
			Presentation Skills	
			Problem Solving	
			Safety	
			Technical Skill	
			Training Performance	
Written Communication				
Getting Along	Interpersonal	Building Relationships	Agreeableness Surgency/Extraversion	
		Communication		
		Consultative Skills		
		Cooperating		
		Influence		
		Interpersonal Skill		
		Organizational Citizenship		
		Service Orientation		
		Teamwork		
	Trustworthiness			
	Intrapersonal		Dependability	Conscientiousness Emotional Stability
			Detail Orientation	
			Flexibility	
			Following Procedures	
			Integrity	
Planning				
Respect				
Risk Taking				
Stress Tolerance				
Work Attitude				

Table 2

PIC/HPI Scale Definitions

Scale Name	Definition
	The degree to which performance improves if an incumbent. . . .
Adjustment	Is calm, self-accepting, and even-tempered.
Ambition	Is socially self-confident and competitive.
Sociability	Seems to need and/or enjoy interacting with others.
Interpersonal Sensitivity	Is seen as perceptive, tactful, and socially sensitive.
Prudence	Seems conscientious, conforming, and dependable.
Inquisitive	Seems creative and interested in intellectual matters.
Learning Approach	Seems to value learning/training for its own sake.

Table 3

Descriptive Statistics for Most Important CET Ratings

Competency	Mean SME Rating	Standard Deviation	Performance Domain
Integrity	3.69	0.50	Intrapersonal
Decision Making	3.50	0.52	Technical
Business Acumen	3.38	0.51	Leadership
Resource Management	3.38	0.57	Leadership
Flexibility	3.28	0.55	Intrapersonal
Stress Tolerance	3.28	0.60	Intrapersonal
Achievement	3.20	0.58	Leadership
Teamwork	3.19	0.61	Interpersonal
Communication	3.17	0.69	Interpersonal
Initiative	3.17	0.57	Leadership
Leadership	3.15	0.54	Leadership
Planning	3.14	0.66	Intrapersonal
Employee Development	3.13	0.61	Leadership

Note. N = 225. Ratings made on a five-point scale, ranging from “0” (*Not associated with job performance*), to “4” (*Critical to job performance*).

Table 4

Meta-Analytic Correlations between HPI Scales and Overall Performance for the Managerial Job Family

HPI Scale	<i>N</i>	<i>K</i>	<i>r_{obs}</i>	ρ_v	ρ
Adjustment	3,751	35	.15	.20	.22
Ambition	3,751	35	.21	.29	.31
Sociability	3,751	35	.05	.07	.08
Interpersonal Sensitivity	3,751	35	.09	.13	.15
Prudence	3,751	35	.08	.11	.13
Inquisitive	3,751	35	.05	.07	.08
Learning Approach	3,074	31	.07	.10	.12

Note. *N* = number of participants across *K* studies; *K* = number of studies; *r_{obs}* = mean observed validity; ρ_v = operational validity (corrected for range restriction and criterion reliability only); ρ = true validity at scale level (corrected for range restriction and predictor-criterion reliability).

Table 5

Correlations Between the HPI Scales and Competencies for Study #324

Competency	ADJ	AMB	SOC	INT	PRU	INQ	LEA
Dependability	.30 (.22)**	.16 (.11)	-.22 (-.16)**	.05 (.03)	.20 (.14)*	-.02 (-.02)	.07 (.05)
Initiative	.34 (.24)**	.36 (.25)**	-.08 (-.06)	.22 (.16)**	.18 (.12)*	.09 (.06)	.22 (.16)**
Stress Tol.	.47 (.34)**	.23 (.16)**	-.16 (-.12)*	.26 (.19)**	.28 (.20)**	.01 (.01)	.14 (.10)
Communication	.36 (.25)**	.19 (.13)*	-.13 (-.09)	.17 (.12)*	.11 (.08)	-.09 (-.07)	.05 (.04)
Flexibility	.26 (.18)**	.10 (.07)	-.19 (-.14)*	.08 (.06)	.16 (.11)	-.07 (-.05)	.03 (.02)
Integrity	.34 (.24)**	.07 (.05)	-.21 (-.15)**	.21 (.15)**	.17 (.12)*	-.02 (-.01)	-.06 (-.05)
Teamwork	.33 (.23)**	.20 (.15)*	-.21 (-.15)*	.14 (.10)	.20 (.14)*	-.03 (-.02)	.09 (.06)
Work Attitude	.36 (.25)**	.23 (.16)**	-.11 (-.08)	.07 (.05)	.13 (.10)	.02 (.01)	.13 (.09)
Respect	.30 (.22)**	.05 (.03)	-.20 (-.14)*	.12 (.08)	.12 (.08)	-.15 (-.11)	-.05 (-.03)
Bus. Acumen	.30 (.21)**	.13 (.09)	-.15 (-.11)	.09 (.06)	.03 (.02)	-.01 (-.01)	-.05 (-.03)
Employ. Dev.	.39 (.28)**	.36 (.26)**	-.10 (-.07)	.16 (.11)	.16 (.12)*	.07 (.05)	.18 (.13)*
Leadership	.35 (.25)**	.30 (.21)**	-.12 (-.08)	.13 (.09)	.22 (.16)**	.10 (.07)	.15 (.11)
Average	.34 (.24)	.20 (.14)	-.16 (-.11)	.14 (.10)	.16 (.12)	-.01 (-.01)	.08 (.05)

Note. Correlations are corrected for criterion unreliability. Uncorrected correlations appear in parentheses. $N = 290$; * $p \leq .05$,

one-tailed. ** $p \leq .01$, one-tailed. ADJ = Adjustment; AMB = Ambition; SOC = Sociability; INT = Interpersonal Sensitivity; PRU = Prudence; INQ = Inquisitive; LEA = Learning Approach.

Table 6

Meta-Analysis Results for Synthetic/Job Component Validities from Hogan Archives

Competency / HPI Scale	<i>N</i>	<i>K</i>	<i>r_{obs}</i>	ρ_v	ρ
Integrity					
Adjustment	3660	36	.11	.16	.17
Ambition	3660	36	.02	.02	.02
Sociability	3660	36	-.03	-.04	-.05
Interpersonal Sensitivity	3660	36	.08	.11	.13
Prudence	3660	36	.15	.21	.24
Inquisitive	3660	36	-.02	-.03	-.03
Learning Approach	3520	34	.02	.03	.03
Decision Making					
Adjustment	1105	8	.08	.12	.12
Ambition	1105	8	.13	.19	.20
Sociability	1105	8	.07	.10	.11
Interpersonal Sensitivity	1105	8	.04	.05	.06
Prudence	1105	8	.00	-.01	-.01
Inquisitive	1105	8	.13	.18	.20
Learning Approach	1105	8	.09	.13	.15
Business Acumen					
Adjustment	89	1	.31	.43	.46
Ambition	89	1	.34	.48	.51
Sociability	89	1	.06	.09	.10
Interpersonal Sensitivity	89	1	.18	.25	.30
Prudence	89	1	.11	.15	.17
Inquisitive	89	1	.15	.22	.25
Learning Approach	--	--	--	--	--

Competency / HPI Scale	<i>N</i>	<i>K</i>	<i>r_{obs}</i>	ρ_v	ρ
Resource Management					
Adjustment	381	3	-.11	-.15	-.16
Ambition	381	3	.21	.30	.32
Sociability	381	3	.22	.30	.33
Interpersonal Sensitivity	381	3	.00	.00	.00
Prudence	381	3	-.04	-.05	-.06
Inquisitive	381	3	.16	.22	.25
Learning Approach	381	3	-.02	-.03	-.03
Flexibility					
Adjustment	3126	22	.11	.16	.17
Ambition	3126	22	.14	.20	.21
Sociability	3126	22	.06	.08	.09
Interpersonal Sensitivity	3126	22	.06	.09	.10
Prudence	3126	22	.04	.05	.06
Inquisitive	3126	22	.05	.07	.08
Learning Approach	3026	20	.06	.08	.09
Stress Tolerance					
Adjustment	5676	52	.20	.28	.30
Ambition	5676	52	.06	.09	.10
Sociability	5676	52	-.03	-.04	-.04
Interpersonal Sensitivity	5676	52	.10	.14	.16
Prudence	5676	52	.12	.17	.19
Inquisitive	5676	52	.00	.00	.00
Learning Approach	5536	50	.03	.05	.05

Competency / HPI Scale	<i>N</i>	<i>K</i>	<i>r_{obs}</i>	ρ_v	ρ
Achievement					
Adjustment	4496	48	.06	.08	.09
Ambition	4496	48	.13	.19	.20
Sociability	4496	48	.00	.00	.00
Interpersonal Sensitivity	4496	48	.02	.02	.03
Prudence	4496	48	.04	.06	.07
Inquisitive	4496	48	.01	.02	.02
Learning Approach	4496	48	.02	.03	.04
Teamwork					
Adjustment	4417	36	.13	.18	.19
Ambition	4417	36	.03	.05	.05
Sociability	4417	36	-.02	-.03	-.04
Interpersonal Sensitivity	4417	36	.08	.11	.13
Prudence	4417	36	.12	.18	.20
Inquisitive	4417	36	-.02	-.03	-.03
Learning Approach	3804	31	.03	.05	.05
Communication					
Adjustment	5225	51	.07	.10	.11
Ambition	5225	51	.08	.12	.13
Sociability	5225	51	.02	.03	.03
Interpersonal Sensitivity	5225	51	.06	.09	.10
Prudence	5225	51	.04	.06	.07
Inquisitive	5225	51	.02	.03	.04
Learning Approach	4971	48	.04	.05	.05

Competency / HPI Scale	<i>N</i>	<i>K</i>	<i>r_{obs}</i>	ρ_v	ρ
Initiative					
Adjustment	3947	26	.11	.16	.17
Ambition	3947	26	.21	.29	.32
Sociability	3947	26	.06	.09	.09
Interpersonal Sensitivity	3947	26	.05	.07	.08
Prudence	3947	26	.08	.12	.13
Inquisitive	3947	26	.09	.13	.14
Learning Approach	3557	22	.09	.12	.13
Leadership					
Adjustment	3205	24	.10	.14	.15
Ambition	3205	24	.19	.27	.29
Sociability	3205	24	.08	.11	.12
Interpersonal Sensitivity	3205	24	.02	.03	.04
Prudence	3205	24	.05	.07	.08
Inquisitive	3205	24	.06	.09	.10
Learning Approach	3205	24	.04	.06	.07
Planning					
Adjustment	2166	22	.07	.10	.11
Ambition	2166	22	.09	.13	.14
Sociability	2166	22	.01	.01	.01
Interpersonal Sensitivity	2166	22	.04	.05	.06
Prudence	2166	22	.09	.13	.14
Inquisitive	2166	22	-.01	-.01	-.01
Learning Approach	1893	19	.03	.03	.04

Competency / HPI Scale	N	K	r_{obs}	ρ_v	ρ
Employee Development					
Adjustment	1414	10	.04	.06	.06
Ambition	1414	10	.19	.27	.29
Sociability	1414	10	.10	.14	.16
Interpersonal Sensitivity	1414	10	.08	.11	.14
Prudence	1414	10	.06	.08	.09
Inquisitive	1414	10	.06	.08	.10
Learning Approach	1364	9	.02	.03	.03

Note. N = number of participants across K studies; K = number of studies; r_{obs} = mean observed validity; ρ_v = operational validity (corrected for range restriction and criterion reliability only); ρ = true validity at scale level (corrected for range restriction and predictor-criterion reliability).

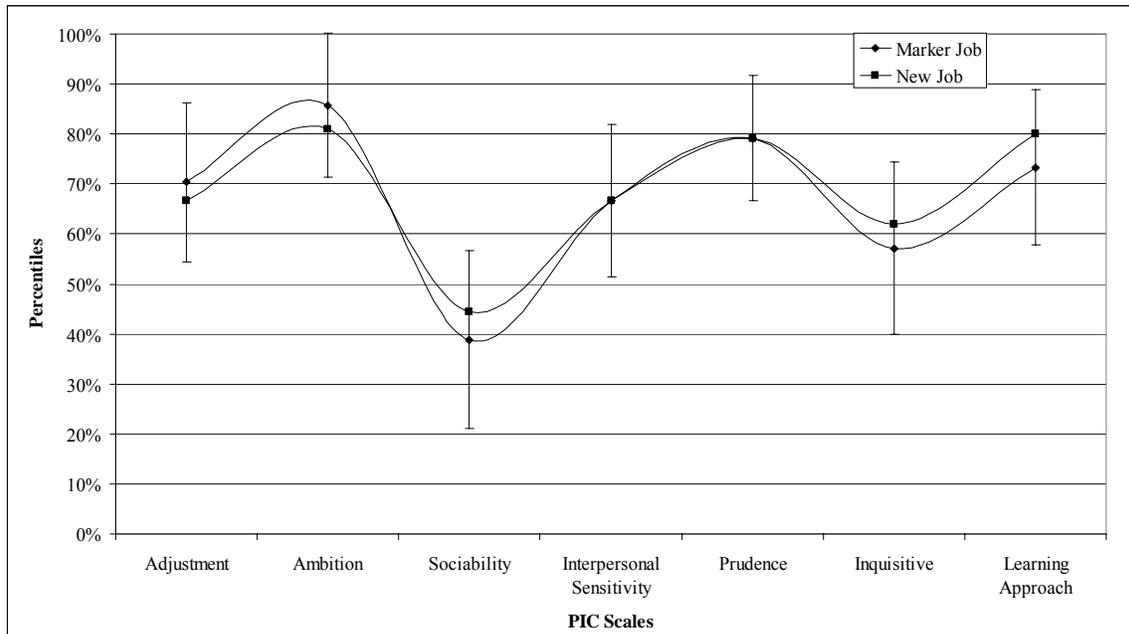
Table 7

Combined Validity Generalization Results for the Managerial Job

	Adjustment	Ambition	Interpersonal Sensitivity	Prudence	Learning Approach	Total Validity for Predictor Battery
Meta-analysis	.22	.31	.15	.13	.12	.42
Transportability	.34	.20	.14	.16	.08	.41
Synthetic	.15	.21	.10	.11	.06	.28
Total Validity by Predictor	.23	.24	.13	.13	.09	.37

Note. Validity coefficients are summarized in the text. Marginal column entries for predictors across methods are the means of the three methods. Marginal row entries for methods across predictors are the means for the synthesized personality test batteries using Nunnally (1978).

Figure Caption

Figure 1. Comparison of PIC Profiles for Marker and New Managerial Jobs

Appendix

Sample Items from the JET

JET Section

PIC

The 48-item PIC job analysis identifies (a) the personal characteristics needed to execute successfully the requirements of a job, and (b) the degree to which possession of these personal characteristics improves job performance

Sample Items

Would job performance IMPROVE if a _____ *...?

- Is steady under pressure
 - Has clear career goals
 - Seems to enjoy social interaction
 - Is kind and considerate
-

DCQ

The 22-item DCQ job analysis identifies (a) personal characteristics that can inhibit performance in a specified job, and (b) the degree to which these personal characteristics degrade job performance

Sample Items

Would job performance DECLINE if a _____ *...?

- Becomes irritable when frustrated
 - Resents criticism and takes it personally
 - Avoids taking any risks
 - Is typically silent and uncommunicative
-

MIC

The 40-item MIC job analysis assesses the environment in which an employee works and the values that help define workgroup climate

Sample Items

The _____ * work group(s) in our organization...

- Do things to improve the appearance of offices and facilities
 - Enjoy meeting new people
 - Show sympathy for those with personal problems
 - Set clear financial goals for the work group
-

CET

The CET allows subject matter experts to indicate the criticality of performance competencies to successful job performance

Sample Items

- Stress Tolerance--Handles pressure without getting upset, moody, or anxious
 - Teamwork--Works well in groups and is a good team player
 - Decision Making--Evaluates issues and uses sound reasoning to make decisions
 - Initiative--Takes action before being told what to do
-

Cognitive JA

The 30-item cognitive job analysis measure allows subject matter expert to assess the strategic reasoning, tactical reasoning, and operations for reasoning components of job performance

Sample Items

Please rate the importance of each reasoning statement for performance on the job.

- Anticipate opportunities that will affect planning and direction
 - Follow problems back to their cause
 - Interpret information in text, numeric, or graphic media accurately
-

*Blanks represent where job titles should be inserted.