TECHNICAL BRIEF

Hogan Research Methodology

Chapter 3: Validity Generalization Studies

- Meta-Analysis
  - The Five-Factor Model and Job Performance
  - Personality-Based Validity Coefficient Benchmarking
  - Summary of Meta-Analysis Results for Generalizing Validity of Five-Factor Model Personality Measures
  - Gathering Meta-Analysis Evidence for Generalizing Validity of the HPI and HDS at the Job Family Level
- Transportability of Validity
  - Gathering Transportability Validity Evidence
- Synthetic/Job Component Validity
  - Gathering Synthetic Validity Evidence

This technical brief is an excerpt of the Validity Generalization chapter from the Hogan Research Methodology (HRM) addendum for all technical documentation.

A3. VALIDITY GENERALIZATION STUDIES

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 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 was due to statistical and procedural factors (Guion, 1998, p. 368) — idiosyncratic factors that could be ignored or statistically corrected.

Many psychologists now agree that "validity" is a unitary concept, not a type of method or an attribute of a test. Guion and Highouse (2006, p. 134) define validity as "a property of the inferences drawn from test scores." In addition, many psychologists now agree that more ways exist to assess the validity of inferences from test scores than a specific local study of their relationship with job relevant criteria (McPhail, 2007). When available, researchers may use Validity Generalization (hereafter, VG) evidence 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*:

> 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)

Given this guidance from the *Principles*, Hogan employs all three methods for establishing VG evidence: (1) meta-analysis, (2) transportability of validity, and (3) synthetic/job component validity (J. Hogan, Davies et al., 2007). The following sections describe each VG method in more detail and outlines the conditions in which Hogan uses them in the research process.

A3.1. Meta-Analysis

Schmidt and Hunter (1977) introduced meta-analysis to psychometric research, which 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 (Hunter, 1980; Schmidt & Hunter, 1977), biographical data inventories (Schmidt & Rothstein, 1994), personality inventories (Barrick & Mount, 1991; Barrick et al., 2003; Dudley et al., 2006; J. Hogan & Holland, 2003; Judge, Bono et al., 2002; Salgado, 1997, 1998; Tett, Jackson, & Rothstein, 1991), assessment center exercises (Arthur, Day, McNelly, & Edens, 2003; Meriac, Hoffman, Woehr, & Fleisher, 2008), and situational judgment tests (McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001) generalize across studies.

The *Principles* recognize the usefulness of meta-analysis and Aguinis and Pierce (1988) state that meta-analysis is 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." (p. 28)

Of the three VG methods, meta-analysis provides the most generalizable results, but 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. J. Hogan, Davies et al. (2007) demonstrate the use of all three methods in combination.

Meta-analysis averages findings from multiple studies of the same relationship to provide a best estimate of ρ (i.e., the population correlation) by controlling for error due to sampling, measurement range restriction, and unreliability in predictor and criterion measures (Smith & Glass, 1977). In addition, meta-analyses include carefully developed criteria for deciding which 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 researchers are less likely to publish studies with insignificant results. Rosenthal (1979) notes that such omissions are problematic for meta-analysis research as they produce results based on too 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. J. Hogan and Holland (2003) did the same using a domain skills model as the basis for a meta-analysis of the validity of personality predictors. 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.

A3.1a. The Five-Factor Model and Job Performance

Table A1 presents the results of six large-scale meta-analyses summarizing relations between the FFM scales and overall job performance. Note that the correlations presented in the table are *uncorrected* estimates. Across studies, the Conscientiousness/Prudence scale appears to be the most consistent predictor of job performance. The Emotional Stability/Adjustment and Agreeableness/Interpersonal Sensitivity scales also predict performance across studies, although the correlation coefficients are generally smaller than those of the Conscientiousness/Prudence scale.

Table A1 FFM Meta-Analysis Results: Uncorrected Validity Estimates

	FFM Scales						
Study	ADJ	AMB	SOC	INP	PRU	INQ	LRN
A.	.15	.10	.10	.22	.12	.18	.18
B.	.05	.01	.01	.04	.12	.01	.01
C.	.09	.05	.05	.01	.10	.04	.04
D.	.09	.06	.06	.07	.14	.04	.04
E.	.25	.20	NA	.18	.22	.20	.15
F.	.17	.22	.22	.06	.20	.16	.16

Note. ADJ = Emotional Stability/Adjustment; AMB = Ambition/Extraversion; SOC = Extraversion/Sociability; INP = Interpersonal Sensitivity/Agreeableness; PRU = Conscientiousness/Prudence; INQ = Openness/Inquisitive; LRN = Openness/Learning Approach. A = Tett et al. (1991). Sample sizes = 280 (Agreeableness) to 2,302 (Extraversion). B = Barrick & Mount (1991). Sample sizes = 3,694 (Emotional Stability) to 4,588 (Conscientiousness). C = Salgado (1997). Sample sizes = 2,722 (Openness) to 3,877 (Emotional Stability). D = Hurtz & Donovan (2000). Sample sizes = 5,525 (Openness) to 8,083 (Conscientiousness). E = J. Hogan & Holland (2003). Sample sizes = 1,190 (Inquisitive) to 3,698 (Ambition). F = Judge, Bono, et al., (2002). Sample sizes = 7,221 (Openness) to 11,705 (Extraversion). NA = Not Available.

Unlike earlier meta-analyses, which evaluated the validity of the FFM in relation to indices of overall performance, J. Hogan and Holland (2003) aligned the FFM scales with performance criteria. Prompted by earlier calls for research to align predictors with criterion, (Ashton, 1998; J. Hogan & Roberts, 1996; Paunonen, Rothstein, & Jackson, 1999), J. Hogan and Holland meta-analyzed 43 independent samples (*N* = 5,242) that contained HPI and criterion data. For this analysis, J. Hogan and Holland aligned HPI scales with criterion measures reflecting FFM themes. As seen in Table A2, the relations between HPI scales and aligned performance ratings proved stronger than previous FFM research. Results indicated the following operational validities: Adjustment = .37, Ambition = .31, Interpersonal Sensitivity = .25, Prudence = .31, Inquisitive = .29, Learning Approach = .22. The fully corrected correlation coefficients ranged from .25 (HPI Learning Approach) to .43 (HPI Adjustment).

Table A2 Meta-Analysis Results for HPI Scales with Construct-Aligned Criteria

HPI Scale	*N*	*K*	r_{obs}	ρ_v	ρ
Adjustment	2,573	24	.25	.37	.43
Ambition	3,698	28	.20	.31	.35
Sociability	N/A	N/A	N/A	N/A	N/A
Interpersonal Sensitivity	2,500	17	.18	.25	.34
Prudence	3,379	26	.22	.31	.36
Inquisitive	1,190	7	.20	.29	.34
Learning Approach	1,366	9	.15	.22	.25

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 unreliability; ρ = true validity at scale level corrected for range restriction and predictor-criterion unreliability; N/A indicates insufficient data to compute meta-analysis. All observed correlations are statistically significant at $p < .05$.

In application, organizations should use multiple personality scales to screen job applicants. Using multiple scales accounts for the various personal characteristics necessary for success, as any one scale is unlikely to map the entire performance domain of any job. J. Hogan and Holland (2003) also illustrate the value of using multiple scales. For example, to predict criteria concerning the ability to tolerate stress, the HPI Adjustment scale is the best single predictor. However, to predict resourceful problem solving or the ability to generate creative solutions, the HPI Inquisitive scale yields the largest validity coefficient. In addition to using multiple personality scales to predict performance, Schmidt and Hunter (1998) provided evidence supporting incremental validity of personality measures over General Mental Ability (GMA), or "g." In reviewing over 85 years of selection research, Schmidt and Hunter showed that adding a measure of Conscientiousness to GMA tests improved validity by 18%. Furthermore, the addition of an integrity measure to GMA improved validity by 27%, the largest increment across 18 other selection measures (e.g., work sample tests, interviews, job knowledge, biographical data, and assessment centers).

Across these studies, the meta-analysis results support the generalizability of the Conscientiousness / Prudence, Emotional Stability / Adjustment, and Agreeableness / Interpersonal Sensitivity measures to multiple occupations and industries. Moreover, the results from J. Hogan and Holland (2003) support the generalizability of every scale on the HPI except Sociability for predicting personality-saturated criteria. Empirical evidence supports validity generalization of three FFM measures (Conscientiousness, Emotional Stability, and Agreeableness) in general, and six of the seven HPI scales in particular.

A3.1b. Personality-Based Validity Coefficient Benchmarking

Criteria used to designate a "meaningful" predictor-criterion correlation remain poorly defined. Consequently, researchers define the meaningfulness of a correlation solely on its magnitude, which is reasonable but not sufficient. Interpreting the usefulness of a correlation coefficient based solely on magnitude is one strategy, since the percentage of variance accounted for in the criterion increases with the magnitude of the correlation. However, at what point does the magnitude of a correlation become "meaningful"? Is it .10, .20, .30, or

.70? Rather than focus exclusively on the magnitude of observed correlation coefficients, a benchmarking strategy is more appropriate.

The assessment literature includes many studies that evaluate the validity of the FFM personality measures across jobs, organizations, and industry types. Hough and Oswald (2008) summarize some of the major findings. These studies reflect the appropriate benchmark from which to evaluate the validity of the FFM scales. By comparing validity coefficients found in this technical report to the validity coefficients reported in the peer-reviewed literature, it is possible to derive some general conclusions about the validity and utility of potential personality predictors of job performance.

To establish a benchmark from which to compare the generalized validity coefficients presented in this report, Table A3 summarizes the sample-weighted validity coefficients of various predictors reported in the scientific literature. The sample-weighted validity of GMA tests, which are widely regarded as the "best" predictors of job performance, is only r = .21. Relative to the sample-weighted validity coefficients reported by J. Hogan and Holland (2003), the validity of GMA appears less predictive of construct-oriented criteria (not overall supervisory ratings of job performance) than the HPI Adjustment and Prudence scales.

Table A3 Comparative Validity of Assessments for Predicting Overall Job Performance

Study	Predictor	r_{obs}
A.	Conscientiousness Tests	.18
B.	Integrity Tests	.21
C.	Structured Interviews	.18
D.	Unstructured Interviews	.11
E.	Situational Judgment Tests	.20
F.	Biodata	.22
G.	General Mental Ability	.21
H.	Assessment Centers	.28
I.	Resumes	.18

Note. r_{obs} = mean observed validity; A = Mount & Barrick (2001). B = Ones et al. (1993). C & D = McDaniel, Whetzel, Schmidt, & Maurer (1994). E = McDaniel, Hartman, Whetzel, & Grubb (2007). F = Bliesener (1996). G = Pearlman, Schmidt, & Hunter (1980). H = Arthur et al. (2003). I = O'Leary (2009).

Also noteworthy are the validity coefficients of FFM scales reported in five other meta-analyses (see Table A9). Excluding J. Hogan and Holland's (2003) results, the validity of Emotional Stability measures ranges between .05 (Barrick & Mount, 1991) and .17 (Judge, Bono, et al., 2002). A similar pattern exists for Conscientiousness measures, with validity coefficients ranging between .10 (Salgado, 1997) and .20 (Judge, Bono et al., 2002). For the remaining FFM scales, only Tett et al. (1991) and Judge, Bono, et al. (2002) report validity coefficients at or above .10.

J. Hogan and Holland (2003) present validity coefficients (see Table A9) that are, on average, 24% larger in magnitude than the highest correlation coefficients reported in previous personality-based meta-analyses. There are three important differences between the J.

Hogan and Holland study and previous meta-analyses. First, they aligned predictors with indices of job performance. J. Hogan and Holland reasoned that personality scales are not designed to be omnibus predictors of job performance, but rather to predict *facets* of job performance. By matching predictors and performance criteria, the observed validities increased. Campbell (1990) articulated this construct alignment strategy, although it is seldom used. Second, most early studies evaluating the validity of FFM personality scales relied on classification schemes to translate scales from non-FFM instruments (e.g., California Psychological Inventory; Gough, 1987) into the FFM domains. During the classification process, raters misclassified scales into FFM dimensions. When errors like this occur, validity decreases. Finally, J. Hogan and Holland relied on a single personality tool (HPI), which eliminated the possibility of coding or classification errors. Together these three factors help untangle the personality literature and establish the appropriate benchmark from which to evaluate the validity of personality scales in occupational settings.

A3.1c. Summary of Meta-Analysis Results for Generalizing Validity of Five-Factor Model Personality Measures

Researchers are skeptical about the merits of some procedures used in meta-analyses. In particular, they believe corrections can be used inappropriately to overestimate predictor-criterion relationships. Nonetheless, the meta-analyses described above provide lower bound estimates of the validity of personality measures for predicting job performance. Reviewing the research on meta-analysis evidence permits certain conclusions. First, meta-analysis results strongly support the validity of Conscientiousness measures for predicting various job criteria, including overall job performance. Second, evidence to support the generalized validity of Emotional Stability and Agreeableness for job performance is moderate to strong, particularly as the criterion becomes more saturated with requirements for interpersonal skill(s). Lastly, the validity coefficients for Extraversion/Surgency measures (particularly the HPI Ambition scale) are strong for predicting criteria associated with achieving results and leading others. The remaining Five Factor dimension — Intellect/Openness to Experience — is not as generalizable as the others because it is relevant for a smaller range of jobs and criteria.

A3.1d. Gathering Meta-Analysis Evidence for Generalizing Validity of the HPI and HDS at the Job Family Level

When conducting a VG study, Hogan uses a meta-analysis procedure to identify HPI and HDS scales important to job performance at the job family level. The Hogan archive contains hundreds of studies examining jobs classified into one of seven job families (see Table A4 for a detailed description of each Hogan job family). As stated in the HPI technical manual (R. Hogan & J. Hogan, 2007):

> Job families are groups of occupations classified as similar based on work performed, skills, education, training and credentials required for competence. The seven job families used for this analysis were derived from nine "job classifications" used by the Equal Employment Opportunity Commission (EEO) for employers in the United States. These nine EEO classifications are used to capture information about an organization's

ethnic make-up. We used this scheme for two reasons: (a) a large percentage of employers within the United States are familiar with EEO job classifications; and (b) the job classifications are conceptually clear and easy to use for reporting purposes. (p. 79)

Table A4 Hogan Job Family Definitions

Hogan Job Family	Definition
Managers & Executives	Employees assigned to positions of administrative or managerial authority over the human, physical, and financial resources of the organization.
Professionals	Employees with little legitimate authority, but high status within the organization because of the knowledge and/or skills they possess. These employees are usually experts with a broad educational background and rely primarily on their knowledge and intellect to perform their duties.
Technicians & Specialists	Employees who rely on the application of highly specific knowledge in skilled manipulation (e.g., operation, repair, cleaning, and/or preparation) of specialized technology, tools, and/or machinery.
Sales & Customer Support	Employees who use appropriate interpersonal style and communication techniques to establish relationships, sell products or services that fulfill customers' needs and provide courteous and helpful service to customers after the sale.
Administrative & Clerical	Employees who plan, direct, or coordinate supportive services of an organization. The main function of these employees is to facilitate the function of professionals by completing jobs that require little formal education or skill to complete (e.g., professional assistants, secretaries, and clerks).
Operations & Trades	Employees who are craft workers (skilled), operatives (semi-skilled), and laborers (unskilled) whose job knowledge and skills are primarily gained through on-the-job training and experience; little prerequisite knowledge or skill is needed.
Service & Support	Employees that perform protective services for individuals and communities (e.g., police, fire fighters, guards) and non-protective services for individuals that require little to no formal training but a high degree of interaction with people (e.g., food service, recreation and amusement).

Using job analysis information (e.g., job descriptions, focus group information, O*NET codes), multiple Hogan researchers classify a job into the appropriate job family and then identify relevant studies from the Hogan archive. Based on studies within each job family, we meta-analyze validity coefficients for each HPI and HDS scale.

Hogan uses the procedures specified by Hunter and Schmidt (1990) to accumulate results across studies and assess effect sizes. All studies use zero-order product-moment correlations, which eliminates the need to convert alternative statistics to values of *r*. We report operational validities, which we correct for sampling error, unreliability in the criterion measure, and range restriction. We do not correct correlation coefficients for predictor unreliability to estimate validity at the construct level. Although some (e.g., Mount & Barrick, 1995; Ones et al., 1993) argue this is a relevant artifact that can be corrected, Hogan believes

it is premature to estimate the validity of a perfect construct when there is no firm agreement on the definition of the construct itself. Results, therefore, represent relationships between HPI scales and job performance.

Hunter and Schmidt (1990) argue that different samples should contribute the same number of correlations to meta-analysis results to avoid bias. Thus, Hogan selected only one correlation per study so that each sample contributed only one point estimate per predictor scale. Hogan also computed a range restriction index for HPI scales. Following procedures described by Hunter and Schmidt (1990), Hogan divides each HPI scale's within-study standard deviation by the standard deviation reported by R. Hogan and J. Hogan (1995). This procedure produces an index of range restriction for each HPI scale for each study. We use mean replacement within job family to estimate range restriction correction factors when within study standard deviation is unavailable.

Although some researchers (e.g., Murphy & De Shon, 2000) argue against the use of rater-based reliability estimates, Hogan follows procedures outlined by Barrick and Mount (1991) and Tett et al. (1991), and uses the .52 reliability coefficient proposed by Tett et al. to estimate the reliability of supervisory ratings of job performance.

Note that meta-analysis evidence for the validity of the MVPI is unavailable because the MVPI is not a generalizable predictor of job performance, since workplace culture and motivators are not consistent across companies or even specific job families (Lock & Bourdreau, 2004).

A3.2. Transportability of Validity

The next step in the VG process involves transporting validity evidence established for one job and using it as a foundation for candidate screening in a similar job. The *Uniform Guidelines* supports transportability of validity and is the primary reference for determining when it is appropriate to transport validity evidence from one job to another. In addition, Hoffman, McPhail, and colleagues (Hoffman & McPhail, 1998; Tippins, McPhail, Hoffman, & Gibson, 1999) discuss the technical requirements that should be satisfied before transporting validity evidence in situations that preclude local validation. Finally, 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 capable of justifying the appropriateness and applicability of a selection procedure. This assumes that the original validation study is technically sound and the target and referent jobs can be described as "closely related" (*Bernard v. Gulf Oil Corp.,* 1981). Situations where transportability might apply include those in which organizations must choose a selection procedure for the same job across multiple locations and different companies or for different jobs with similar requirements. It might also be a useful strategy for validating screening guidelines for different job titles within a single job family (see Gibson & Caplinger, 2007).

The *Uniform Guidelines*, the *Standards*, and the *Principles* all recognize transportability of selection procedures (cf. Tippins, 2003). Although employment discrimination experts distinguish between these three documents, Hogan focuses on their common themes. For

example, all three require that the original research be technically adequate. The *Uniform Guidelines* emphasize the need for evidence regarding fairness, validity, and job similarity as criteria for transportability. Personality-based selection procedures typically yield no adverse impact, satisfying requirements set by the *Uniform Guidelines* and precedents set in many courts (Lindemann & Grossman, 1996). However, it should be noted fairness is considered a social rather than a psychometric issue.

The *Standards* emphasize the need for good cumulative research (e.g., meta-analysis) and discourage reliance on a single local validation study as a foundation for transportability of validity unless the referent study is "exceptionally sound." Interestingly, the original design for transportability of a selection procedure relies on a single referent validation study. The *Principles* emphasize the importance of establishing similarity between the original (referent) and target jobs. Researchers can establish evidence of similarity based on job requirements, job context, and job applicants. For personality-based selection systems, demonstrating job similarity has been challenging because few personality-related job analysis methods were available. Notable exceptions are Raymark, Schmit, and Guion (1997) and Hogan's JET methodology (Foster et al., 2012). Hogan estimates similarity using converging evidence and professional judgment.

A3.2a. Gathering Transportability Validity Evidence

Hogan conducts transportability of validity research by analyzing the current target job in qualitative and quantitative terms. First, Hogan identifies a marker job for which a criterion-related validity study already exists in the Hogan archive. Next, Hogan establishes similarity between the target and marker jobs through close alignment of job descriptions, O*NET codes, and JET profiles. The O*NET typology provides a standard external metric for rating job similarity.

Hogan uses JET profiles (i.e., PIC for HPI, DCQ for HDS) to evaluate the similarity of the target job to the marker job in the Hogan archive. We determine similarity by calculating the standard error of the means (Mitchell & Jolley, 2010; SE_{mean}) for each of the relevant JET scales for the target job using the following formula:

$$SE_{mean} = \frac{S}{\sqrt{N}}$$

Where S is the standard deviation of the JET scale SME ratings and N is the number of SMEs that complete the JET.

Next, we construct 95% confidence intervals for each scale by adding and subtracting 1.96 SE_{mean} to and from each raw score scale mean for the target job. We then compare each PIC scale mean for the marker job with the confidence interval for the same scale for the target job. The jobs are sufficiently similar for transportability of validity if most of the seven scale means for the marker job fall within the 95% confidence intervals of the target job.

In cases where Hogan cannot identify a specific job in the Hogan archive that meets the stringent requirements of single-study transportability, Hogan identifies multiple similar jobs

(i.e., very similar in respect to the tasks and responsibilities associated with performing the job, but not close enough for single-study transportability). Hogan nominates jobs in the archive that are similar enough to the target job to be used as a hybrid form of transportability validity evidence. Instead of using the results from a single study, Hogan meta-analyzes the results across all relevant studies to provide an alternative form of transportability of validity evidence. This procedure may also be used when multiple studies meeting the stringent requirements for single study transportability exist in the Hogan archive.

Some jobs are too complex to meet the stringent requirements necessary for transportability of validity. In these instances, Hogan defers to the meta-analysis and the synthetic/job component validation sections.

A3.3. Synthetic/Job Component Validity

The most specific validity generalizability evidence results from synthetic validity/job component validity research. Mossholder and Arvey (1984) noted that, where meta-analysis relies on global evaluations of job similarity, synthetic validity requires a more detailed examination of the work. The strategy is criterion driven and involves finding the best set of predictors comprehensively representative of the criterion space.

Lawshe (1952) introduced synthetic validity over 50 years ago. With a few notable exceptions (e.g., Guion, 1965; McCormick, DeNisi, & Shaw, 1979; Primoff, 1959), early researchers largely ignored the approach because they believed that assessment validity was specific to situations. The interpretive review and demonstration by Mossholder and Arvey (1984) is a rare exception. Mossholder and Arvey defined 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 prior criterion-related studies predicting those components. We then "synthesize" the valid predictors of the key job components into an assessment battery for the new job (Balma, 1959; Lawshe, 1952). Since Mossholder and Arvey's initial demonstration, synthetic validity has gained more support and popularity (e.g. Hoffman, Holden, & Gale, 2000; Jeanneret & Strong, 2003; Johnson & Carter, 2010; Johnson, Carter, Davison, & Oliver, 2001; Johnson et al., 2010; McCloy, 1994; 2001; Scherbaum, 2005).

Brannick and Levine (2002) point out that synthetic validity approaches allow us to build up validity evidence from small samples with common job components. Johnson and Carter (2010) showed that synthetic validity (a) produced coefficients quite similar to coefficients obtained from more traditional local validation research and (b) may be more advantageous when developing selection batteries for newly created jobs, given that tenured job incumbents are needed for criterion-related validation studies.

The *Uniform Guidelines* are vague about technical requirements and documentation for synthetic/job component validity, but the *Principles* explicitly include this strategy. Synthetic validation involves (a) identifying the important components of a job or jobs composing a job family, (b) reviewing prior research on the prediction of each component, and (c) aggregating correlations across multiple studies for each component of the job to form a test battery

(Scherbaum, 2005). Mossholder and Arvey (1984) summarized 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 validational strategy. (p. 323)

Job Component Validity (hereafter, JCV: McCormick et al., 1979) is one type of synthetic validity. Jeanneret (1992) described JCV as falling "within the rubric of construct validity" (p. 84). Researchers have primarily used JCV to study the cognitive demands of jobs by correlating job dimensions using Position Analysis Questionnaire data (Jeanneret, 1992; Hoffman, Rashkovsky, & D'Egidio, 2007). 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. Because the concept of synthetic validity has evolved over the years, Hogan uses interchangeably the terms criteria, performance dimensions, job components, work components, competencies, and domains of work.

A3.3a. Gathering Synthetic Validity Evidence

The first step in synthetic validation is conducting a job analysis to determine the important components of the job. This involves identifying the most highly rated competencies across subject matter experts using the CET section of the JET. Next, Hogan maps these CET items to the updated Hogan Competency Model (HCM; Hogan Assessment Systems, 2012). A crosswalk between CET items and the updated competency model is displayed in table A5. Then, Hogan identifies predictor(s) for each competency deemed important for job performance. The Hogan archive provides a means to identify the best predictor(s) of each competency. Foster, Lemming, and Johnson (2010) mapped each of the criteria from over 250 criterion-related validity studies in the Hogan archive onto the HCM and conducted a meta-analysis for each scale-by-competency relationship. These meta-analyses provide stable estimates of the relationships between the 7 HPI scales, the 11 HDS scales, and the critical competencies as rated by SMEs. They report operational validities, which they corrected for sampling error, unreliability in the criterion measure, dichotomization (when necessary), and range restriction.

Table A5 Crosswalk between Competency Labels in CET and the Hogan Competency Model

CET Label	HCM Label	HCM Definition
Achievement Orientation	Driving for Results	Accomplishes goals, completes tasks, and achieves results.
Active Listening	Listening to Others	Listens and restates the ideas and opinions of others to improve mutual understanding.
Ambiguity Tolerance	Dealing with Ambiguity	Comfortably handles unclear or unpredictable situations.
Building Relationships	Relationship Building	Develops collaborative relationships to facilitate current and future objectives.
Building Teams	Team Building	Assembles productive groups based upon required skills, goals and tasks.
Business Acumen	Business Insight	Applies business knowledge to achieve organizational goals and objectives.
Caring	Caring about People	Displays sensitivity towards the attitudes, feelings, or circumstances of others.
Citizenship	Organizational Citizenship	Exceeds job requirements to help the organization.
Competitive	Competing with Others	Strives to exceed others' performance.
Decision Making	Decision Making	Uses sound judgment to make timely and effective decisions.
Delegation	Delegating	Assigns work to others based on tasks, skills, and workloads.
Dependability	Dependability	Performs work in a reliable, consistent, and timely manner.
Detail Orientation	Detail Focus	Performs work with care, accuracy, and attention to detail.
Employee Development	Developing People	Provides support, coaching, training, and career direction to others.
Financial Acumen	Financial Insight	Applies financial knowledge to achieve organizational goals and objectives.
Flexibility	Flexibility	Changes direction as appropriate based on new ideas, approaches, and strategies.
Following Procedures	Rule Compliance	Adheres to directions, policies, and/or legal guidelines.
Goal Setting	Setting Goals	Identifies short-term objectives and steps to achieve them.
Industry Knowledge	Industry Insight	Applies knowledge of industry trends and outlooks to achieve organizational goals and objectives.
Influence	Influencing Others	Persuades others to help achieve organizational goals and objectives.
Information Analysis	Processing Information	Gathers, organizes, and analyzes diverse sources of information.
Initiative	Taking Initiative	Takes action without needing direction from others.
Innovation	Driving Innovation	Stimulates creative ideas and perspectives that add value.
Interpersonal Skills	Leveraging People Skills	Gets along well with others, is tactful, and behaves appropriately in social situations.
Intrapersonal Skills	Self Management	Demonstrates appropriate motivation, attitude, and self-control.
Leadership	Leading Others	Demonstrates general leadership ability and effectiveness.
Managing Change	Driving Change	Champions new methods, systems, and processes to improve performance.
Managing Conflict	Managing Conflict	Resolves hostilities and disagreements between others.
Managing Performance	Driving Performance	Provides guidance and feedback to maximize performance of individuals and/or groups.
Motivating Others	Inspiring Others	Motivates others to accomplish organizational goals.
Negotiation	Negotiating	Explores alternatives to reach outcomes acceptable to all parties.
Oral Communication	Verbal Communication	Expresses ideas and opinions effectively in spoken conversations.
Organizational Commitment	Engagement	Demonstrates loyalty and commitment through enthusiasm and extra effort.
Perseverance	Overcoming Obstacles	Pursues goals and strategies despite discouragement or opposition.
Planning/Organizing	Planning and Organizing	Coordinates and directs activities to help achieve business objectives.

Table A13 Crosswalk between Competency Labels in CET and the Hogan Competency Model

CET Label	HCM Label	HCM Definition
Political Awareness	Political Savvy	Recognizes, interprets, and works within the political environment of an organization.
Presentation Skills	Presenting to Others	Conveys ideas and information to groups.
Problem Identification	Anticipating Problems	Forecasts and detects errors, gaps, and potential flaws.
Problem Solving	Solving Problems	Identifies solutions given available information.
Professionalism	Professionalism	Acts in accordance with job-related values, principles, and standards.
Quality Orientation	Quality Focus	Strives to meet quality standards and produce quality work products.
Resource Management	Managing Resources	Coordinates people and financial and material capital to maximize efficiency and performance.
Responsibility	Accountability	Accepts responsibility for one's actions regardless of outcomes.
Risk Management	Taking Smart Risks	Evaluates tradeoffs between potential costs and benefits and acts accordingly.
Safety	Safety Focus	Attends to precautions and proper procedures to guard against work-related accidents and injuries.
Sales Ability	Sales Focus	Generates revenue by promoting products and services to others.
Self Confidence	Displaying Confidence	Projects poise and self-assurance when completing work tasks.
Self Development	Self Development	Actively acquires new knowledge and skills to remain current with and/or grow beyond job requirements.
Service Orientation	Customer Focus	Provides courteous, timely, and helpful service to encourage client loyalty.
Social Engagement	Networking	Builds and maintains a system of strategic business connections.
Strategic Planning	Driving Strategy	Directs effort to achieve long-term business objectives.
Stress Tolerance	Handling Stress	Manages pressure without getting upset, moody, or anxious.
Talent Management	Attracting Talent	Recruits, rewards, and retains individuals with needed skills and abilities.
Teamwork	Teamwork	Collaborates with others to achieve goals.
Time Management	Time Management	Plans and prioritizes work to maximize efficiency and minimize downtime.
Trustworthiness	Integrity	Acts honestly in accordance with moral or ethical principles.
Valuing Diversity	Leveraging Diversity	Respects and values individual differences to obtain a desired effect or result.
Vigilance	Staying Alert	Remains focused when performing monotonous tasks.
Work Attitude	Positive Attitude	Displays a positive disposition towards work.
Work Ethic	Working Hard	Consistently strives to complete tasks and assignments at work.
Work Skills	Leveraging Work Skills	Applies technology and job-relevant abilities to complete work tasks.
Written Communication	Written Communication	Expresses ideas and opinions effectively in writing.

REFERENCES

Aguinis, H., & Pierce, C. A. (1998). Testing moderator variable hypotheses meta-analytically. *Journal of Management, 24,* 577–592.

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.

Ashton, M. C. (1998). Personality and job performance: The importance of narrow traits. *Journal of Organizational Behavior, 19,* 289-303.

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.

Bernard v. Gulf Oil Co., 619 F.2d 495, 470 (5th Cir. 1980), aff _d, 452 U.S. 89 (1981).

Bliesener, T. (1996). Methodological moderators in validating biographical data in personnel selection. *Journal of Occupational and Organizational Psychology, 69*, 107-120.

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* (pp. 265-294). Thousand Oaks, CA: Sage.

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 1, 2nd ed., pp. 39-74). Palo Alto, CA: Consulting Psychologists Press.

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.

Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A meta-analytic investigation of Conscientiousness in the prediction of job performance: Examining the intercorrelations and the incremental validity of narrow traits. *Journal of Applied Psychology, 91*, 40-57.

Foster, J., Gaddis, B., & Hogan, J. (2012). Personality-based job analysis. In M. A. Wilson, W. Bennett, S. G. Gibson, & G. M. Alliger (Eds.), *The handbook of work analysis: Methods, systems, applications, and science of work measurement in organizations* (pp. 247-264). New York, NY: Routledge Taylor & Francis Group.

Foster, J., Lemming, M., & Johnson, A. (2010). *Validity of the Hogan Personality Inventory for competencies and job family profiles*. Tulsa, OK: Hogan Press.

Gatewood, R. D., & Feild, H. S. (1994). *Human resource selection* (3rd Ed.). Orlando, FL: The Dryden Press.

Ghiselli, E. E. (1966). *The validity of occupational aptitude tests*. New York, NY : Wiley.

Ghiselli, E. E., & Brown, C. H. (1955). *Personnel and industrial psychology* (2nd ed.). New York, NY: McGraw-Hill.

Gibson, W. M., & Caplinger, J. A. (2007) Transportation of validation risks. In S. M. McPhail (Ed.), *Alternative validation strategies* (pp. 29-81). San Francisco, CA: Jossey-Bass.

Gough, H. G. (1987) *Manual for the California Psychological Inventory.* Palo Alto, CA: Consulting Psychologists Press.

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.

Guion, R. M., & Highhouse, S. (2006). *Essentials of personnel assessment and selection.* Mahwah, NJ: Lawrence Erlbaum.

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.

Hoffman, C. C., Rashkovsky, B., & D'Egidio, E. (2007). Job component validity: Background, current research, and applications. In S. M. McPhail (Ed.), *Alternative validation strategies* (pp. 82-121). San Francisco, CA: Jossey-Bass.

Hogan Assessment Systems. (2012). *The development of the Hogan Competency Model and competency-based predictive algorithms*. Tulsa, OK: Hogan Press.

Hogan, J., Davies, S., & Hogan, R. (2007). Generalizing personality-based validity evidence. In S. M. McPhail (Ed.), *Alternative validation strategies* (pp. 181-229). San Francisco, CA: Jossey-Bass.

Hogan, J., & Hogan, R. (1996). *Motives, Values, Preferences Inventory manual.* Tulsa, OK: Hogan Assessment Systems.

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, R., & Hogan, J. (1995). *Hogan Personality Inventory manual* (2nd ed.). Tulsa, OK: Hogan Assessment Systems.

Hogan, R., & Hogan, J. (2007). *Hogan Personality Inventory manual* (3rd ed.). Tulsa, OK: Hogan Assessment Systems.

Hough, L. M., & Oswald, F. L. (2008). Personality testing and industrial-organizational psychology: Reflections, progress, and prospects. *Industrial Organizational Psychologist, 1,* 272-290.

Hunter, J. E. (1980). *Validity generalization for 12,000 jobs: An application of synthetic validity and validity generalization to the General Aptitude Test Battery (GATB)*. Washington, DC: U.S. Department of Labor, Employment Service.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis.* Newbury Park, CA: Sage.

Jeanneret, P. R. (1992). Applications of job component/synthetic validity to construct validity. *Human Performance, 5,* 81-96.

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.

Johnson, J. W., & Carter, G. W. (2010). Validating synthetic validation: Comparing traditional and synthetic validity coefficients. *Personnel Psychology, 63*, 755-795.

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., Steel, P., Scherbaum, C. A., Hoffman, C. C., Jeanneret, P. R., & Foster, J. (2010). Validation is like motor oil: Synthetic is better. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 3*, 305-328.

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.

Lawshe, C. H. (1952). What can industrial psychology do for small business? (A symposium). *Personnel Psychology, 5,* 31-34.

Lindemann, B., & Grossman, P. (1996). *Employment discrimination law* (3rd ed.). Washington, DC: American Bar Association.

Lock, J. D., & Boudreau, N. (2004, April). *Same job different values: Comparing similar jobs across organizations*. Paper presented at the 19th Annual Conference of the Society for Industrial and Organizational Psychology, Chicago, IL.

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., 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.

McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. L. (2007). Situational judgment test, response instructions, and validity: A meta-analysis. *Personnel Psychology, 60*, 63-91.

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, 84,* 730-740.

McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology, 79,* 599-616.

McPhail, S. M. (2007). *Alternative validation strategies*. San Francisco, CA: Jossey-Bass.

Mecham, R. C. (1985, August). Comparative effectiveness of situational, generalized, and job component validation methods. In P. R. Jeanneret (Chair), *Job component validity: Job requirements, estimates, and validity generalization comparisons.* Symposium conducted at the 92nd Annual Convention of the American Psychological Association, Los Angeles, CA.

Meriac, J. P., Hoffman, B. J., Woehr, D. J., & Fleisher, M. S. (2008). Further evidence for the validity of assessment center dimensions: A meta-analysis of the incremental criterion-related validity of dimension ratings. *Journal of Applied Psychology, 93,* 1042-1052.

Mitchell, M. L., & Jolley, J. M. (2010). *Research design explained* (7th ed.). Belmont, CA: Wadsworth.

Mossholder, K. W., & Arvey, R. D. (1984). Synthetic validity: A conceptual and comparative review. *Journal of Applied Psychology, 69,* 322-333.

Mount, M. K., & Barrick, M. R. (1995). The Big Five personality dimensions: Implications for research and practice in human resources management. *Research in Personnel and Human Resources Management, 13,* 153-200.

Mount, M. K., & Barrick, M. R. (2001). *Personal Characteristics Inventory manual.* Libertyville, IL: Wonderlic.

Murphy, K. R., & De Shon, R. (2000). Interrater correlations do not estimate the reliability of job performance ratings. *Personnel Psychology, 53,* 873-900.

O'Leary, B. (2009). *Position paper regarding the potential and drawbacks of resumes in the federal hiring system* (Center for Talent Services). Washington, DC: U.S. Office of Personnel Management.

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.

Paunonen, S. V., Rothstein, M. G., & Jackson, D. N. (1999). Narrow reasoning about the use of broad measures for personnel selection. *Journal of Organizational Behavior, 17,* 639-655.

Pearlman, K., Schmidt, F. L, & Hunter, J. E. (1980). Validity generalization results for tests used to predict job proficiency and training success in clerical occupations. *Journal of Applied Psychology, 65,* 373-406.

Primoff, E. S. (1959) Empirical validation of the J-coefficient. *Personnel Psychology, 12,* 413-418.

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

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*, 638-641.

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,* 529-54.

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

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.

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.

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.

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., 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.

