

Are You Interested? A Meta-Analysis of Relations Between Vocational Interests and Employee Performance and Turnover

Chad H. Van Iddekinge
Florida State University

Philip L. Roth
Clemson University

Dan J. Putka
Human Resources Research Organization, Alexandria, Virginia

Stephen E. Lanivich
Florida State University

A common belief among researchers is that vocational interests have limited value for personnel selection. However, no comprehensive quantitative summaries of interests validity research have been conducted to substantiate claims for or against the use of interests. To help address this gap, we conducted a meta-analysis of relations between interests and employee performance and turnover using data from 74 studies and 141 independent samples. Overall validity estimates (corrected for measurement error in the criterion but not for range restriction) for single interest scales were .14 for job performance, .26 for training performance, -.19 for turnover intentions, and -.15 for actual turnover. Several factors appeared to moderate interest–criterion relations. For example, validity estimates were larger when interests were theoretically relevant to the work performed in the target job. The type of interest scale also moderated validity, such that corrected validities were larger for scales designed to assess interests relevant to a particular job or vocation (e.g., .23 for job performance) than for scales designed to assess a single, job-relevant realistic, investigative, artistic, social, enterprising, or conventional (i.e., RIASEC) interest (.10) or a basic interest (.11). Finally, validity estimates were largest when studies used multiple interests for prediction, either by using a single job or vocation focused scale (which tend to tap multiple interests) or by using a regression-weighted composite of several RIASEC or basic interest scales. Overall, the results suggest that vocational interests may hold more promise for predicting employee performance and turnover than researchers may have thought.

Keywords: vocational interests, personnel selection, criterion-related validity, job performance, turnover

Selection researchers long have been concerned with the identification of constructs that predict criteria such as job performance. Researchers have identified a range of useful predictor constructs, including mental ability, personality, motivation, and knowledge and skills (e.g., J. P. Campbell, McCloy, Oppler, & Sager, 1993; Schmidt, Hunter, & Outerbridge, 1986). Researchers also have studied the validity of various selection procedures designed to assess these constructs, including tests, background data questionnaires, structured interviews, work samples, and assessment centers (e.g., Schmidt & Hunter, 1998).

One set of predictor constructs that has received comparatively little attention in the selection literature is vocational interests, which are traits that reflect preferences for certain types of work activities and environments (Harmon, Hansen, Borgen, & Hammer, 1994). Although interests have a long and established history within the vocational psychology literature (e.g., Holland, 1959; Kuder, 1939; Strong, 1943), very little recent research has addressed the use of interests for selection. For example, more recent reviews of selection research have not even mentioned interests (e.g., Ployhart, 2006; Sackett & Lievens, 2008; Schmitt, Cortina, Ingerick, & Wiechmann, 2003).

Furthermore, unlike other selection procedures (e.g., cognitive ability tests, personality assessments, structured interviews, assessment centers), no comprehensive meta-analyses have summarized the validity evidence for vocational interests with respect to criteria most relevant to selection researchers. Instead, existing knowledge about the potential value of interests for selection appears to be based primarily on a few small-scale quantitative summaries. The only published meta-analysis we know of in this area is Hunter and Hunter (1984). These researchers meta-analyzed relations between scores on the Strong Interest Inventory (Strong, 1943) and various criteria. Mean validity estimates were .10 for supervisor ratings of job performance ($k = 3$), .25 for job level attained ($k = 4$), .18 for training success ($k = 2$), and .22 for tenure ($k = 3$). With the exception of the supervisor ratings validity

This article was published Online First July 11, 2011.

Chad H. Van Iddekinge and Stephen E. Lanivich, Department of Management, Florida State University; Philip L. Roth, Department of Management, Clemson University; Dan J. Putka, Human Resources Research Organization, Alexandria, Virginia.

An earlier version of this article was presented at the 71st Annual Meeting of the Academy of Management, San Antonio, Texas, August 2011. We thank John Hausknecht, In-Sue Oh, and Fred Oswald for their helpful comments on an earlier version of the article. We also are grateful to Norman (Trip) Wyckoff for helping obtain unpublished studies for possible inclusion in the meta-analysis.

Correspondence concerning this article should be addressed to Chad H. Van Iddekinge, College of Business, Florida State University, 821 Academic Way, P.O. Box 3061110, Tallahassee, FL 32306-1110. E-mail: cvanidde@fsu.edu

estimate, which was corrected for measurement error in the criterion, these values appear to represent observed (i.e., uncorrected) estimates of validity.

We also found two unpublished studies that summarized validity evidence for interests with respect to employee performance. First, Barge and Hough (1988) reported mean observed validities between interests and various measures of job and training performance. For job performance, median observed validity estimates were .20 for performance ratings ($k = 11$) and .33 for productivity records ($k = 3$). For training performance, median observed validity estimates were .17, .35, and .28 ($k = 8, 2$, and 3) for training grades, instructor ratings, and course completion, respectively. No corrected validity estimates were reported. Second, Morris (2003) reported a mean observed correlation of .24 between interests and job performance ($k = 27$). This correlation increased to .29 when corrected for unreliability in both the predictor and criterion (a validity estimate corrected for measurement error in the criterion only was not reported).

Although these studies provide a useful starting point for understanding how interests may relate to employee performance at work, there is much more we need to know. For example, conclusions from Hunter and Hunter (1984) and Barge and Hough (1988) are based on very small sets of primary studies, and thus it is difficult to know how stable and generalizable their validity estimates may be. Although Morris's (2003) dissertation research was based on relatively more studies, he only considered studies that used some form of congruence index to measure interests. As we discuss later, interests can be measured in various ways, and congruence indices may represent a less than optimal way to use scores from multiple interest scales for prediction. Further, all of these studies appeared to include some primary studies whose designs and measures may lead to ambiguous conclusions regarding the criterion-related validity of interests. For example, both Barge and Hough and Morris included studies that reported statistically significant results only, or used atypical job performance criteria such as salary, managerial level, and self-ratings of performance (Hunter & Hunter, 1984, also appeared to include some of these criteria). In addition, some of the training performance validities included in Barge and Hough's work study were based on student samples and criteria that reflected academic performance (e.g., college grade point average).

The goal of the present article is to present an updated meta-analysis of the criterion-related validity of vocational interests. The results of this study extend the limited prior work in this area in three main ways. First, we cumulate a much larger set of primary studies (i.e., 74 studies and 141 independent samples) than the previous quantitative summaries did. Second, we estimate relations between interests and a wider range of outcomes organizations may value; namely, job performance, training performance, and turnover, including both turnover intentions and actual turnover. Lastly, we investigate several potential moderators of interest-criterion relations. For example, we examine whether the job relevance of interest constructs moderates validity, and whether some types of interest scales are better predictors of criteria than others (e.g., Holland's, 1973, interest constructs vs. job-specific interests). In addition, we examine whether study features such as sample (applicants vs. employees), design (predictive vs. concurrent), and setting (civilian vs. military) function as moderators.

We begin by providing a brief overview of vocational interests and Holland's (1973, 1997) theory. We then describe theoretical bases for why and how interests may influence performance and turnover. Finally, we discuss several potential moderators of interest-criterion relations.

Vocational Interests

Vocational interests reflect relatively stable individual differences that influence behavior through preferences for certain work activities and work environments (Holland, 1973; Kuder, 1977; Mount, Barrick, Scullen, & Rounds, 2005). Three aspects of this definition are noteworthy (Van Iddekinge, Putka, & Campbell, 2011). First, within persons, interests tend to be quite stable over time (e.g., Low, Yoon, Roberts, & Rounds, 2005) and, thus, are considered to have a strong dispositional component (although experiences can help shape people's interests; Lent, Brown, & Hackett, 1994). Second, interests are rooted in the work context and focus on the types of activities people prefer to perform and the environments in which they prefer to perform those activities. Third, interests are thought to influence the way people behave at work by increasing motivation to perform the work activities they prefer and by inspiring workers to increase knowledge and skills relevant to performing those activities.

Initial research on interests can be traced back to the work of Strong and Kuder in the 1920s and 1930s (e.g., Kuder, 1939; Strong, 1943). Since then, Holland's (1973, 1997) six-factor model has become the best known and most widely researched interests theory. According to Holland, interests can be used to categorize individuals into six types: realistic, investigative, artistic, social, enterprising, and conventional (i.e., RIASEC). Individuals with realistic interests like practical, hands-on types of work activities. Investigative types like scholarly, intellectual, and scientific types of work. Artistic types like work that involves creative, expressive, and unconventional activities. Social types like work that involves helping, teaching, and caring for others. Enterprising types like work that involves assertive, persuasive, and leadership-oriented activities. Finally, conventional types like work that involves well-ordered and routine activities. Although Holland's theory is not without its detractors (e.g., Gati, 1991; Tinsley, 2000), the RIASEC model is fairly widely accepted in the interests literature.

Theoretical Basis for the Predictive Validity of Vocational Interests

Interests and Employee Performance

Several theoretical perspectives propose or imply that interests are related to job performance. For example, although Holland's (1973, 1997) theory focuses on how congruence between interests and work environment affects vocational choice, it also predicts that congruence will influence satisfaction, "achievement," and continuance in that occupation. Researchers have suggested that interests affect employee performance in two main ways. The first is through motivation (e.g., J. P. Campbell et al., 1993; Hogan & Blake, 1996; Lent et al., 1994; Low & Rounds, 2005; Mount et al., 2005). For instance, social cognitive career theory (Lent et al., 1994) suggests that interests affect (a) people's career-related goals, (b) the actions

they take to achieve those goals, and (c) their performance with respect to those goals.

J. P. Campbell et al.'s (1993) theory on the antecedents and determinants of individual differences in job performance offers similar predictions. Specifically, this theory proposes that interests influence choices people make about (a) which work activities to pursue (or to avoid), (b) how much effort to exert toward those activities, and (c) how long to persist in those activities. Consider the job of a human resources (HR) manager, for example. This job often involves leading people and making decisions, working with and providing service to others, and following procedures and working with data. On any given day, employees in this job must balance these and other performance requirements and make choices about where to invest their effort, how much effort to invest, and how long to persist with that effort. We suggest that employees will tend to gravitate toward, and be more motivated to perform, types of work they find most interesting. Thus, HR managers who possess high levels of enterprising, social, and conventional interests—which are reflected in the performance requirements listed above—may be more motivated to perform than HR managers who possess low levels of such interests.

A second way researchers have suggested that interests affect performance is through the acquisition of job-relevant knowledge and skills (e.g., Ackerman, 1996; Barrick, Mount, & Gupta, 2003; Lent et al., 1994; Sullivan & Hansen, 2004). For example, Ackerman's (1996) PPIK (intelligence-as-process, personality, interests, and intelligence-as-knowledge) theory suggests that although intelligence is related to knowledge development, "the motivational aspects of interests will determine the orientation toward (or away from) the various knowledge domains" (p. 245). In an employment context, individuals whose interests are congruent with the knowledge and skill requirements of a given job may be more motivated to acquire relevant knowledge and skills, which, in turn, may allow them to perform better than individuals whose interests are not congruent with the job's requirements. For example, if a job involves primarily realistic tasks (e.g., as opposed to artistic tasks), new employees who have high realistic interests may be more motivated to learn those tasks than employees who have low realistic interests.

The idea that interests affect motivation, knowledge, and skills also would seem to suggest a possible relationship between interests and how employees perform during training. For example, interests may influence training performance through trainee's motivation to learn, which is thought to be a direct determinant of training outcomes (Noe & Colquitt, 2002). Specifically, interests may affect choices employees make about whether to attend training, how much effort they exert during training, and how long they persist in training. To the extent training reflects the knowledge and skill requirements of the job, employees whose interests are congruent with the job may be more motivated to perform during training (and, in turn, acquire job-relevant knowledge and skills) than people whose interests are not as congruent with the job.

Interests and Turnover

Several related theoretical perspectives provide a basis for a relationship between vocational interests and turnover. According to Holland's (1973, 1997) congruence theory, people's work behavior is partially determined by the congruence between their

interests and the extent to which the environment supports those interests. People whose interests are congruent with a given work environment will (a) be more likely to enter that environment, (b) be more satisfied in that environment, and (c) choose to stay in that environment longer than people whose interests are incongruent with that environment.

Holland's (1973, 1997) work represents one of the theories that laid the foundation for contemporary theory on person-environment (PE) fit. PE fit is broadly defined as the congruence, match, or similarity between the person and the environment (Edwards, 2008). The basic prediction of the theory is that the match between people and their work environments will relate to attitudes such as job satisfaction and organizational commitment, which, in turn, will relate to behaviors such as absenteeism, intentions to turnover, and actual turnover. The specific type of PE fit that is most applicable to vocational interests is that of *needs-supplies fit*. Specifically, the interests a particular work environment supplies serves to fulfill what employees need or find important (Dawis & Lofquist, 1984; Edwards & Shipp, 2007). When the work environment supports people's needs in terms of work they like to perform, then a need will be fulfilled. This, in turn, is thought to produce positive job attitudes (e.g., job satisfaction), and people with more positive attitudes are more likely to remain in that work environment than are people who have less positive attitudes.

Schneider's (1987) attraction-selection-attrition (ASA) theory predicts that social forces in organizations create an environment predisposed toward homogeneity with respect to personality but also to competencies and interests (Schneider, Smith, Taylor, & Fleenor, 1998). With regard to interests, the theory suggests that people are attracted to jobs and organizations they think are aligned with their interests. Then, through selection, organizations choose individuals whose interests are consistent with the goals of the organization. Finally, employees are posited to stay in (i.e., do not attrit from) jobs that support their interests. However, when an employee realizes there is a mismatch between his or her interests and the work he or she must perform, or the environment in which the work is performed, the employee may decide to leave the job or organization.

The above theories provide a basis to predict a relationship between vocational interests and turnover. As a key precursor to actual turnover (Griffeth, Hom, & Gaertner, 2000; Tett & Meyer, 1993), researchers often use turnover intentions as more proximal outcome. Thus, in addition to turnover, we cumulate evidence concerning relations between interests and turnover intentions.

Proposed Moderators

In the sections above, we offer several reasons why vocational interests, in general, may be related to various criteria of concern to selection researchers. However, interests are somewhat unique in how they have been operationalized, measured, and linked to criteria. In this section, we consider how such factors may moderate relations between interests and criteria.

Job Relevance of Interests

Constructs such as general mental ability and conscientiousness are positively associated with performance in almost all types of

jobs (e.g., Barrick & Mount, 1991; Hunter & Hunter, 1984). In contrast, the magnitude and direction of relations between individual interest constructs and criteria likely will vary by job. For example, there may be positive relations between enterprising interests and outcomes (e.g., job performance) in jobs that support such interests, and nil or perhaps negative relations between enterprising interests and outcomes in jobs that do not support such interests.

We expect job relevance to moderate interest–criterion relations, such that the more theoretically relevant a given interest construct is to the work performed in a given job (i.e., the more an interest is supported by the job), the more scores on interest measures will relate *positively* to the criteria. We also expect the converse to be true. Specifically, the less theoretically relevant a given interest construct is to the work performed in a given job (i.e., the more an interest is *not* supported by the job), the more scores on interest measures will relate *negatively* to the criteria. In other words, to the extent that an interest is clearly supported by the job, or is clearly not supported by the job, we expect validities to increase and be positive in the first case and negative in the second case.

Hypothesis 1: Job relevance will moderate the criterion-related validity of vocational interests. Specifically, the relationship between job relevance and validity will be positive, such that as the relevance of an interest to the job increases, validity coefficients will become more positive, and vice versa.

Type of Interest Scale

Researchers have developed and studied various ways to scale individuals' vocational interests. First, there are *construct focused* scales that assess individuals' standing on the RIASEC dimensions, or interests that can be linked to the RIASEC model. Examples of measures that yield construct focused scales include the Self-Directed Search (Holland, 1994), the Vocational Preference Inventory (Holland, 1985), and the Kuder Preference Record (Kuder, 1939). Second, *basic interest* scales yield scores that reflect interest in a specific type of work or activity but that are not directly linked to the RIASEC dimensions. Basic interests also tend to be somewhat narrower in scope (Day & Rounds, 1997). For example, the Army Vocational Interest Career Examination (AVOICE; Hough, Barge, & Kamp, 2001) yields scales that measure interests in different types of work, such as work involving administrative activities, computers, or vehicle operation.

Third, *job focused* interest scales are developed to predict criteria (e.g., job performance) for a specific job within a specific organization. Such scales comprise items from company-developed or commercially available interest inventories that prior research has shown to differentiate employees with respect to job performance, turnover, and so forth. Similarly, *vocation focused* scales are designed to measure interests relevant to work performed in a particular vocation. For example, the Strong Vocational Interest Blank (Strong, 1943) yields scores for a large number of vocations, such as insurance agents, librarians, and police officers. Unlike construct focused and basic interest scales, which measure single interest dimensions, job and vocation focused scales tend to measure a mix of interests that differentiate members of the target job or vocation.

We expect type of interest scale to moderate criterion-related validity, such that job and vocation focused scales will yield larger validity estimates than construct focused and basic interest scales. For one, job and vocation focused measures assess interests that are closely aligned with the type of work performed in a particular job or vocation. For instance, job focused measures tend to be developed specifically to differentiate successful performers (or stayers) from unsuccessful performers (or leavers) for a given job. In contrast, construct focused and basic interest scales are not tailored to assess interests specific to a particular job or vocation. Furthermore, whereas construct focused and basic interest scales focus on singular interest constructs (e.g., realistic interests) or aspects of work (e.g., clerical activities), job and vocation focused measures are more heterogeneous in that they tend to assess multiple interests that serve to differentiate individuals on that particular job. As such, job and vocation focused scales may be better predictors of criteria such as job performance, which also tend to be heterogeneous.

Hypothesis 2: Type of interest scale will moderate the criterion-related validity of vocational interests, such that validity estimates will be larger for job and vocation focused interest scales than for construct focused and basic interest scales.

Approach to Linking Multiple Interests to Criteria

Relations between interests and criteria are thought to be maximized by considering an individual's interests profile (e.g., profile across the RIASEC dimensions), rather than one's standing on any single interest dimension (e.g., Ployhart, Schneider, & Schmitt, 2006). As discussed, job and vocation focused interest scales often assess multiple interests that are relevant to a particular job, but yield a single score. Another option is to use multiple construct focused or basic interest scales and to combine their scores empirically.

The most common strategy researchers have used to link multiple interest scales to criteria is profile similarity (e.g., D , D^2) or congruence indices (e.g., Iachan index; Iachan, 1984). Unfortunately, congruence indices impose constraints that often are not supported by the data and, therefore, can result in suboptimal prediction of criteria (Edwards & Parry, 1993). In fact, it has been suggested that congruence indices will almost always explain less variance in outcomes than the basic data from which they are calculated (Tinsley, 2000). Congruence indices also are less than optimal from a theoretical perspective because they can mask potentially important differential relations between components of the index (e.g., individual RIASEC dimensions) and the criteria (Cronbach & Gleser, 1953; Edwards, 2002). For instance, in the HR manager example we discussed earlier, social interests may be positively related to performance with respect to working with others but may have no relationship to performance with respect to following procedures. Conversely, conventional interests may be unrelated to working with others but may be positively related to following procedures. Finally, artistic interests may be unrelated to working with others but may be *negatively* related to following procedures. If we used only congruence indices to link interests to criteria, such theoretically and practically important relations could be overlooked.

As an alternative to congruence indices, Edwards has shown how polynomial regression models can be used to relax constraints imposed on relations between predictors (e.g., interests, values) and criteria and, therefore, overcome some of the limitations above (e.g., Edwards & Parry, 1993; Edwards & Shipp, 2007). Although we were not able to examine polynomial regression models in this meta-analysis (i.e., because such analyses would have required access to the raw data from each primary study), we were able to examine how the validity of simple linear regression composites of interests compare to the validity of congruence indices. In addition to the possibility of yielding larger validity estimates, a simple regression approach enables researchers to interpret more clearly the direction and magnitude of each interest–criterion relationship conditional on other interests in the profile. This, in turn, can contribute to both understanding and decision making, such as which interest dimensions to retain and how to weight them for use in selection.

Hypothesis 3: The method for linking multiple interests to criteria will moderate the criterion-related validity of vocational interests, such that validity coefficients will be larger using regression-weighted composites than when using congruence indices.

Additional Potential Moderators

We also explore the potential influence of several additional moderator variables. We investigate the nature of the criterion as a moderator by examining whether estimates of criterion-related validity differ for criteria that reflect ratings of employee performance versus objective measures of performance. We also compare validities for voluntary turnover versus involuntary turnover. In addition, we examine the role of study sample (applicants vs. incumbents), study design (predictive vs. concurrent), study setting (civilian vs. military), and publication status (published vs. unpublished). Because we do not have strong expectations regarding any of these potential moderators, we examine them in an exploratory manner to see whether they may affect the pattern of validity evidence for vocational interests.

Method

Literature Search

We began by searching electronic databases, including PsycINFO, ABI/INFORM Global, and Dissertation Abstracts, for potentially relevant articles. We searched for various combinations of key terms that appear in article titles or abstracts (e.g., “interests and performance or success or productivity or turnover or withdrawal or retention”). We also searched the Defense Technical Information Center (DTIC), which is an online database of research sponsored by the U.S. Department of Defense and other government agencies. In addition to key word searches, we performed separate searches within each database for all articles that mentioned each of 35 interest measures we identified through our research. We also searched electronic and hard copies of programs from the past 10 annual conventions of the Academy of Management and Society for Industrial and Organizational Psychology to identify articles that have not appeared elsewhere in the literature.

Finally, we reviewed the reference sections of the articles we obtained to identify additional articles.

Inclusion Criteria

We used six inclusion criteria to evaluate each primary study. First, we only included studies that reported validity results based on individual-level data. We excluded studies that examined relations at the unit-level of analysis (e.g., Comrey & High, 1955), because relations among aggregate data are not necessarily the same as relations among the underlying individual data (Ostroff, 1993).

Second, we only included studies that measured interests using an inventory completed by participants themselves. We excluded studies in which participants’ interests were inferred from their college major (e.g., Richards, 1993), or that measured congruence between participants’ interests and the interests of their coworkers or supervisor (e.g., Meir, Hadas, & Noyfield, 1997). We also excluded studies in which the predictors appeared to comprise both interests and one or more other constructs, such as work values (e.g., Hesketh, McLachlan, & Gardner, 1992).

Third, we only included studies in which the criteria reflected work-related performance or turnover. For example, we excluded studies in which the criteria reflected academic performance of undergraduate students (e.g., Fritzsche, McIntire, & Yost, 2002), doctoral students (e.g., Kahn & Scott, 1997), or students attending a military academy (e.g., Abrahams, Neumann, & Dann, 1969). The performance criteria we considered included job performance and training performance. Job performance was measured using supervisor or peer ratings or some objective criterion (e.g., sales). We excluded studies in which the criterion reflected self-rated job performance (Dik, 2005), salary level (e.g., Arvey & Dewhirst, 1979), or managerial level (e.g., Grimsley & Jarrett, 1973). However, we included two studies (Bieschke, Herbert, & Bard, 1998; Krebs, Smither, & Hurley, 1991) that used self-reported number of journal publications as a criterion, as this represents a somewhat more objective, verifiable measure of performance. Training performance was measured using training examination scores or instructor ratings.

With regard to turnover, company records were used to measure actual turnover, whereas turnover intentions were measured using employee self-reports. We also planned to include other withdrawal related criteria, such as lateness and absenteeism, but there were too few studies to meta-analyze. We attempted to determine whether the criteria reflected voluntary turnover, involuntary turnover, or some combination of the two. We excluded studies that used tenure as a criterion (e.g., T. H. Stone & Athelstan, 1969) because few, if any, of the studies we reviewed controlled for the fact that employees had different hire dates, which can introduce irrelevant variance into the tenure measure.

The fourth criterion required each study to describe an original study and to provide sufficient details concerning the research design, data analysis, and results. For example, we had to exclude several studies (e.g., Bizot & Goldman, 1993; J. T. Campbell, Otis, Liske, & Prien, 1962; Ronan, 1964) for which we could not be reasonably confident about the meaning of the direction of the validity coefficients (e.g., it was unclear whether high scores on the criterion reflected good or poor performance). We also excluded studies because the particulars (e.g., sample, measures,

results) were not adequately described to be reasonably sure of their meaning (e.g., Craig, 1925; Dann & Abrahams, 1977; Super & Roper, 1941). Finally, we had to exclude many studies that did not report statistics (e.g., r , t , M , and SD) needed to derive estimates of criterion-related validity (e.g., Dyer, 1987; Frank & Wilcox, 1978; Thompson, 1947). However, when possible, we first attempted to contact the authors to obtain the necessary information.

Fifth, we only included studies that appeared to be absent of certain methodological issues. For one, we excluded studies in which items in the interest measure were selected based on their relationship with the criterion and did not attempt to cross-validate the measure (e.g., V. W. Stone, 1961). Additionally, we excluded studies that reported statistically significant validity results only (e.g., Dunnette & Kirchner, 1960), as exclusion of nonsignificant results can lead to overestimates of mean validity across studies (McDaniel & Jones, 1988). Furthermore, we excluded studies in which variance in the predictor, criterion, or both was artificially increased. For instance, we excluded studies that employed extreme-groups designs that collected job performance data on a range of employees, but then used data from the best and poorest employees only to validate the interest measure (e.g., Wiggins & Moody, 1983). Such studies were excluded because they can produce inflated effect sizes as well as reduce the representativeness of the sample (i.e., because cases between the extremes are omitted; Butts & Ng, 2009).

Finally, we included only results based on independent samples. We used the method described by Wood (2008) to identify (and

exclude) studies in which a sample appeared to overlap with a sample from another article authored by the same researchers.

We identified 191 studies that appeared to provide data concerning relations between vocational interests and the focal criteria. Of these, 74 studies (38.7%), representing 141 independent samples, met all the criteria. These studies comprised a mix of published journal articles ($k = 41$), dissertations and theses ($k = 17$), technical reports ($k = 14$), and book chapters ($k = 2$). Table 1 shows the number and percentage of studies we had to exclude according to each inclusion criterion. Although studies were excluded for a range of reasons, the four most prevalent were (a) lack of details needed to derive estimates of criterion-related validity, (b) unclear methods and/or results, (c) criteria reflected academic performance rather than work performance, and (d) the reporting of statistically significant results only. In the Appendix, we provide the main codes and input values for all the studies and independent samples included in the meta-analysis.

Coding of Primary Studies

Two of the authors independently coded all the studies. Both were professors with more than 10 years of research experience, including prior experience coding primary studies for meta-analyses. Before analyzing the data, we estimated our level of agreement on the coding of key variables. For the categorical variables, we used Cohen's kappa to assess agreement. The resulting kappa coefficients were as follows: type of publication (1.00), study setting (1.00), type of sample (.94), type of design (.94), and

Table 1
Number and Percentage of Excluded Studies by Inclusion Criterion

Outcome/inclusion criterion	<i>k</i>	%
Total studies identified for possible inclusion	191	
Studies that passed all inclusion criteria	74	38.7
Studies that did not pass one or more inclusion criteria	117	61.3
1. Individual-level data criterion		
Used unit-level data	1	0.9
2. Measured interests criterion		
Inferred interests from college major	1	0.9
Congruence with interests of coworkers or supervisor	2	1.7
Predictor captured both interests and another construct(s)	4	3.4
3. Work-related performance or turnover criterion		
Criterion reflected academic performance	14	12.0
Criterion was self-rating of job performance	7	6.0
Criterion was salary level	4	3.5
Criterion was position level	2	1.7
Criterion was tenure	2	1.7
Criterion was a mix of performance and turnover	1	0.9
4. Reporting of validity results criterion		
Methods and/or results were unclear	25	21.4
Could not extract necessary statistics to estimate validity	28	23.9
5. Methodological issues criterion		
Reported significant results only	13	11.1
Used extreme groups design	8	6.8
Empirical item selection with no cross-validation	1	0.9
6. Independent sample criterion		
Sample overlapped with another study	4	3.4

Note. Percentages reflect the percent of excluded studies ($k = 117$) that were excluded due to each criterion.

type of interest measure (.93). For the remaining variables, we determined the percentage of times the two authors recorded the same value for each independent sample. The level of rater agreement was as follows: sample size (97%), criterion reliability estimate (96%), and validity estimate (98%). Instances of disagreement were resolved by discussion.

Analyses

Observed validities. We implemented Hunter and Schmidt's (2004) psychometric approach to meta-analysis. We began by identifying (and/or computing) the observed validity coefficient(s) within each primary study. Most studies reported zero-order correlations. When other statistics were reported (e.g., *t*, *M*, and *SD*), we converted them to correlations using formulae provided by Hunter and Schmidt.

Studies reported a variety of validity coefficients depending on the nature and number of interest scales and criteria examined. For studies that reported a single validity coefficient between one interest scale and one criterion measure, we used that coefficient in our analyses. For studies that reported validities for multiple interest scales (e.g., scales for each RIASEC dimension), we recorded the interest scale that was theoretically most relevant to the target job or vocation. For most studies, it was clear which scale was most relevant. For example, if participants were insurance salespersons, and one of the interest scales was the Strong Life Insurance scale, we recorded this scale as theoretically most relevant. Similarly, authors that used a measure of the RIASEC dimensions often stated that a particular dimension (e.g., investigative interests) was most relevant to the job. For studies that examined the RIASEC dimensions, but did not clearly indicate which dimension was most relevant to the job, we searched the job on the Occupational Information Network (O*NET) and recorded the interest with the highest score as most relevant. For example, according to O*NET, the interests profile for the job of bank teller is conventional (95), enterprising (61), realistic (28), social (17), investigative (6), and artistic (0). In this case, we recorded conventional as theoretically most relevant. Finally, in a few instances for which we could not confidently determine the most relevant interest scale, we excluded these studies from the pertinent analyses.

Additionally, many studies reported results for regression-based composites or congruence indices that reflected combinations of multiple interest scales. As such, we recorded "composite-level" validity coefficients when available. If the authors did not report one of these composites, but reported intercorrelations among the scale-level predictors and a criterion, we estimated *R* from the correlation matrix. If a study did not report the necessary intercorrelations, we requested them from the authors.

Corrected validities. We then corrected the observed validity estimates for measurement error in the criterion. We did not correct for predictor unreliability, because we wanted to estimate the operational validity of the interest measures, rather than the construct-level relations between interests and the criteria. For job performance criteria, our database included 21 estimates of interrater reliability for supervisor ratings of performance. For each of these studies, we used the reliability estimate the authors reported.

The approach we used for studies that did not report a reliability estimate was as follows. First, for the 21 studies that reported

interrater reliability estimates (i.e., based on ratings of two or more raters), we used the Spearman-Brown formula to estimate what the single-rater reliability would be. The mean single-rater reliability estimate across these studies was .56, which is slightly higher than previous meta-analytic estimates for supervisor ratings of performance (e.g., .52; Viswesvaran, Ones, & Schmidt, 1996). Furthermore, our experience suggested that interrater reliability estimates tend to be smaller in military settings than in civilian settings. Indeed, the mean single-rater estimates within our database were .66 (*k* = 11) for civilian studies and .45 (*k* = 10) for military studies.

Then, for the studies that did not report a reliability estimate, we attempted to determine the number of individuals who rated each participant. If there was only one rater, or if the authors did not report the number of raters, we used one of the mean single-rater reliability estimates we derived—.66 if it was a civilian study and .45 if it was a military study. If there were multiple raters, we used the mean single-rater reliability and Spearman-Brown to derive a reliability estimate that reflected the number of individuals whose ratings contributed to the job performance criterion. For example, if there were two raters, we "inputted" a value of .80 for civilian studies and a value of .62 for military studies.

Three studies reported reliability estimates for objective performance measures, all of which concerned amount of sales. The mean stability coefficient of .77 across these studies is similar to the .80 estimate previous meta-analyses have used for various objective performance measures (e.g., Roth, Huffcutt, & Bobko, 2003). Therefore, we used the .77 value to correct for measurement error in studies that used objective performance criteria.

Three studies reported reliability estimates for training grades. The mean test-retest coefficient across these studies was .80. This reliability estimate is very consistent with reliabilities reported in meta-analyses within the training literature (e.g., McNatt, 2000; Taylor, Russ-Eft, & Chan, 2005). Thus, we used .80 for studies that did not report a reliability estimate for training performance.

For turnover intentions, three studies reported an estimate of internal consistency reliability (mean α = .94). We used these estimates to correct correlations with turnover intentions for criterion unreliability. In contrast, to be consistent with prior research (e.g., Griffeth et al., 2000; Harrison, Newman, & Roth, 2006; Zimmerman, 2008), we did not correct for measurement error in turnover, which is assumed to be measured with little or no error.

However, we did take two steps to standardize the turnover validity estimates that we analyzed. First, studies reported a mix of biserial correlations (which consider the turnover base rate in the sample and estimate what the validity coefficient would be if turnover was a continuous outcome, which it is not) and point-biserial correlations (which are affected by the turnover base rate and are not "corrected" like the biserial correlations). Thus, we first standardized all the validity estimates to reflect observed (i.e., unadjusted) point-biserial correlations and meta-analyzed these validities. We also corrected all the point-biserial correlations to reflect a standard 50–50 split between leavers and stayers (Kemery, Dunlap, & Griffeth, 1988) and then meta-analyzed these validity estimates. This correction controls the potential differences in turnover base rate across studies, which if not corrected, could falsely indicate the existence of moderator effects (Hunter & Schmidt, 2004; Zimmerman, 2008). The mean turnover base rate

across primary studies was .31. One study did not report the turnover base rate, and thus we used the mean value for this study.

Finally, range restriction may be an issue if interests (or their correlates) are used for selection. Unfortunately, not one primary study reported standard deviations for both the sample (e.g., employees) and the relevant population (e.g., job applicants), which are necessary to derive accurate estimates of range restriction. We therefore were unable to correct for this artifact in our main analyses. However, as part of our follow-up analyses, we examine the use of normative standard deviations to estimate what the range restriction-corrected validities may be. As we will describe, these analyses suggested relatively little range restriction in interest scores. Thus, the validity estimates we report below are not likely to be severely biased (downward) due to range restriction.

Results

Overall Evidence for Criterion-Related Validity of Vocational Interests

Table 2 displays the overall meta-analytic results. We present two sets of results for each criterion. First, we present relations between individual interest scales and the criterion (i.e., "scale-level validity estimates"). Again, for studies that used multiple

interests, we used the validity for the theoretically most relevant scale. Second, we present relations between composites of multiple interest scales and the criterion (i.e., "composite-level validity estimates"). These validities reflect regression-weighted composites (i.e., *Rs*) or congruence indices, which we discuss later.

Because the regression-weighted composites maximize interest-criterion relations in the sample in which the regression weights were estimated, the observed *Rs* can overestimate the magnitude of these relations in the population. Thus, we also present "shrunk" *Rs* adjusted for shrinkage using the positive-part Pratt population *R* formula (Shieh, 2008). The resulting values estimate what validity would be if the regression weights were calculated based on the full population. When interpreting these shrunk *Rs*, it is important to note that they are based on regression models that include *all* interest scales examined in each primary study (even those expected to have no relation with the criterion), and the larger the number of predictors in the regression model, the more validities are expected to shrink. In a selection context, researchers may make a priori decisions (without capitalizing on the local data to make this decision) to include a more limited number of scales thought to relate to the criteria, such as those that have demonstrated validity evidence in past research. Therefore, the shrunk values we report may underestimate the potential validity of interest composites.

Table 2

Overall Meta-Analytic Results for the Criterion-Related Validity of Vocational Interests

Criterion/analysis	<i>k</i>	<i>N</i>	<i>r</i>	ρ	<i>SD_ρ</i>	% VE	95% CI	80% CV
Job performance								
Scale-level validity estimates	80	14,522	.11	.14	.12	33.5	.10, .17	-.02, .30
Composite-level validity estimates								
Regression-based composites ^a								
Observed validity ^b	60	10,669	.28	.37	.09			
Shrunk validity ^c	60	10,669	.14	.18	.12			
Congruence indices	8	1,239	.12	.15	.00	100.0	.09, .21	.15, .15
Training performance								
Scale-level validity estimates	17	3,535	.23	.26	.07	50.5	.21, .31	.16, .35
Regression-based composites ^d								
Observed validity	8	891	.32	.36	.00			
Shrunk validity	8	891	.15	.17	.11			
Turnover intentions								
Scale-level validity estimates	22	3,722	-.19	-.19	.07	53.2	-.15, -.24	-.10, -.29
Regression-based composites								
Observed validity	17	2,150	-.31	-.32	.11			
Shrunk validity	17	2,150	-.24	-.24	.13			
Actual turnover								
Scale-level validity estimates	30	14,453	-.14	-.15	.07	31.6	-.12, -.18	-.07, -.23
Composite-level validity estimates								
Regression-based composites								
Observed validity	16	8,526	-.25	-.33	.12			
Without influential case ^e	15	4,025	-.22	-.23	.10			
Shrunk validity	16	8,526	-.21	-.29	.16			
Without influential case	15	4,025	-.13	-.15	.13			
Congruence indices	4	1,232	-.14	-.15	.03	73.9	-.09, -.21	-.11, -.19

Note. *k* = number of validity coefficients from independent samples; *r* = sample-size weighted mean observed validity; ρ = validity estimate corrected for measurement error in the criterion (for job performance, training performance, and turnover intentions) or corrected to a .50 base rate (for turnover); *SD_ρ* = standard deviation of ρ ; % VE = percentage of variance in ρ accounted for by sampling error and measurement error in the criterion; 95% CI = lower and upper bounds of the 95% confidence interval for ρ ; 80% CV = lower and upper bounds of the 80% credibility value for ρ .

^a Validity estimates for the regression-based composites reflect *Rs* rather than *rs*. ^b For reasons noted in the text, we do not present additional meta-analytic statistics for the composite-level validities. ^c Shrunk validities were calculated using the positive-part Pratt formula for estimating population validity (Shieh, 2008). ^d No studies in our database used congruence indices to predict training performance or turnover intentions. ^e See Footnote 4 regarding identification of influential cases.

Job performance. A total of 80 independent samples were available to examine relations between interests and job performance. The overall, sample-size weighted mean observed validity estimate was .11, and the mean operational validity estimate corrected for measurement error in the criterion was .14. However, the standard deviation of corrected validities (SD_p) was .12, which suggests a substantial amount of variance in validities. Consistent with a large SD_p , sampling error and measurement error accounted for only 33.5% of the variance in corrected validities, and the 80% credibility interval (−.02 to .30) was rather wide. The observed and corrected composite-level validity estimates for the regression-weighted composites were .28 and .37 for job performance. The corresponding shrunken R s were .14 and .18.

Training performance. There were 17 independent samples for the relationship between interests and performance during training. The observed and corrected validity estimates for training performance were .23 and .26, respectively. SD_p (.07) was somewhat smaller for this criterion. Sampling error and measurement error accounted for 50.5% of the corrected validities, and the 80% credibility interval was .16 to .35. Observed and corrected validities for the regression-weighted composites were .32 and .36. Adjusting for shrinkage yielded estimates of .15 and .17, respectively.

Turnover. There were 22 independent samples for turnover intentions and 30 independent samples for actual turnover. The validity estimates for these analyses were coded such that negative values indicate that interests are related to lower intentions to turnover and actual turnover. For turnover intentions, both observed and corrected validity estimates were −.19 (SD_p = .07). Statistical artifacts accounted for 53.2% of the variance in corrected validities, and the 80% credibility interval was −.10 to −.29. The observed and corrected validities for the regression-weighted composites were −.31 and −.32, and the shrunken validity estimates were −.24 and −.24.

For actual turnover, the observed validity estimate was −.14, and the corrected validity estimate was −.15 (SD_p = .07).¹ Statistical artifacts accounted for only 31.6% of the variance in corrected validities, and the 80% credibility interval was −.07 to −.23. For studies that included multiple interest scales, the observed and corrected regression-weighted composite validities were −.25 and −.33, respectively.² However, one study emerged as an influential case, and excluding it from the analysis yielded somewhat smaller validity estimates (−.22 and −.23, respectively).³ The observed and corrected shrunken R s were −.21 and −.29 (−.13 and −.15 without the influential case).

Potential Moderators of Relations Between Vocational Interests and the Criteria

Interest constructs and job relevance. Hypothesis 1 predicted that job relevance would moderate estimates of criterion-related validity, such that as the relevance of an interest to the job increases, validity coefficients will become more positive, and vice versa. We tested this hypothesis using studies that reported validity estimates for construct focused interest scales, so that we could determine how relevant each interest construct measured was to the target job. All of these studies measured one or more of the RIASEC dimensions or dimensions linked to the RIASEC model. Furthermore, we limited these analyses to relations between interests and job performance due to the relatively larger number of validity estimates available for this criterion. Finally, to facilitate

interpretation of the results and to conserve space, we used only the scale-level validity estimates (and not the composite-level validity estimates) for this and subsequent moderator analyses.

We first cumulated validity evidence for each RIASEC dimension across studies. These results are presented in Table 3. As we expected, corrected validities for all dimensions were close to zero and quite variable across studies (SD_p = .07 for enterprising to .17 for artistic). These low mean validities and large variance estimates likely reflect the wide variation in relevance of each RIASEC dimension to the work performed in the jobs examined within the primary studies, as we discuss next. Also note that the lower credibility interval estimates for several interests are quite negative (e.g., −.25 for artistic). This suggests that in addition to interests that theoretically relevant to the job, it may be fruitful to consider interests that are theoretically *least relevant* to the job. We revisit this issue in our follow-up analyses below.

To test whether job relevance of the RIASEC dimensions moderates criterion-related validity, we identified studies that examined interest–criterion relations within a single job for which we could obtain scores regarding the relevance of each RIASEC dimension to the work performed. For civilian jobs, we obtained RIASEC information from O*NET. As part of O*NET's Data Collection Program, employees in various occupations rate the extent to which each RIASEC dimension is reflected in the work they perform (further details regarding this program can be found at onet.rti.org). In addition, we were able to obtain such scores for most of the military jobs in our database from research by Ingerick, Diaz, and Putka (2009).

Table 4 displays descriptive statistics for the job relevance scores. The mean scores for realistic, social, enterprising, and conventional interests suggest that on average, these interests were moderately relevant across the primary studies. In contrast, investigative and artistic interests tended to be less relevant. However, the standard deviations indicated wide variability in job relevance across studies for all the interests, particularly for realistic, social, and investigative interests.

Also shown in Table 4 are results of a series of weighted least squares regression analyses in which the criterion-related validity estimates (i.e., for each RIASEC dimension in relation to job performance) were regressed on the job relevance scores. Each study was

¹ Recall that the corrected values for turnover (which are shown in the ρ column in Tables 2 and 5) estimate what validity would be if the turnover rate in each study was 50%.

² Logistic regression often is used to model relations between a set of predictors and a dichotomous outcome such as turnover (Hosmer & Lemeshow, 2000). However, fitting a logistic model requires raw data from the original primary studies. Thus, the composite-level validities analyses for turnover may violate ordinary least squares (OLS) regression assumptions regarding the distribution of residual errors. At the same time, past turnover research has found that OLS and logistic regression analyses tend to produce very similar conclusions (e.g., Barrick & Zimmerman, 2005).

³ To identify potential influential studies, we used a modified version of the sample adjusted meta-analytic deviancy (SAMD) statistic (Beal, Corey, & Dunlap, 2002; Huffcutt & Arthur, 1995) available in *Meta-Analysis Mark XIII*, a Microsoft Excel-based program developed by Piers Steel. If exclusion of a study changed the original corrected validity estimate by 20% or more, we report the results with and without the influential study (e.g., Cortina, 2003).

Table 3
Meta-Analytic Results for the RIASEC Interests and Job Performance

Variable	<i>k</i>	<i>N</i>	<i>r</i>	ρ	<i>SD</i> _{ρ}	% VE	95% CI	80% CV
Realistic	36	4,448	.01	.01	.10	55.9	-.04, .06	-.12, .14
Investigative	36	4,665	.00	.00	.16	33.0	-.06, .06	-.20, .20
Artistic	36	4,448	-.03	-.03	.17	30.6	-.10, .03	-.25, .19
Social	37	4,669	.02	.02	.15	37.0	-.04, .08	-.17, .21
Enterprising	35	4,597	.01	.01	.07	70.6	-.04, .05	-.08, .10
Conventional	36	4,439	.00	.00	.12	49.1	-.06, .05	-.15, .15

Note. RIASEC = realistic, investigative, artistic, social, enterprising, and conventional; *k* = number of validity coefficients from independent samples; *r* = sample-size weighted mean observed validity; ρ = estimated operational validity (corrected for measurement error in the criterion only); *SD* _{ρ} = standard deviation of ρ ; % VE = percentage of variance in ρ accounted for by sampling error and measurement error in the criterion; 95% CI = lower and upper bounds of the 95% confidence interval for ρ ; 80% CV = lower and upper bounds of the 80% credibility value for ρ .

weighted by the inverse of the sampling error variance, such that studies with less sampling error received greater weight than studies with more sampling error (Hedges & Olkin, 1985; Steel & Kammeyer-Mueller, 2002). All the regression coefficients were positive and modest to moderate in size. This provides support for Hypothesis 1 and suggests that as the job relevance of a particular interest construct increases, the relationship between that interest and job performance becomes more positive (and vice versa).

However, it is important to note that validity estimates for construct focused measures tend to be small, even when a particular RIASEC dimension is relevant to the job. To illustrate, we regressed validity estimates for each RIASEC dimension on the corresponding job relevance scores. We then used job relevance scores of -2 SDs, -1 SD, the mean, $+1$ SD, and $+2$ SDs to derive predicted values for the validity estimates (see Table 4). Although the predicted validities are larger at higher levels of job relevance, the magnitude of these values still is quite small. In fact, none of the predicted validities for the most extreme job relevance scores (i.e., 2 SDs above or below the mean) exceeded .19.

Type of interest scale. Hypothesis 2 predicted that job and vocation focused scales would yield larger validity estimates than would the single, most theoretically relevant construct and basic interest scale for a given job. Table 5 displays results for this and subsequent moderator analyses. As shown, validities for job and vocation focused interest scales consistently were larger than validities for construct focused and basic interest scales. For job performance, corrected validity estimates were .23 for job and vocation focused scales, .10 for construct focused scales (which decreased to .07 when an influential case was excluded), and .11 for basic interest scales. For training performance, the corrected validities were .29 for job and vocation focused scales, .15 for construct focused scales, and .23 for basic interest scales. For turnover, the corrected validities were .17 for job and vocation focused scales, .12 for construct focused scales, and .13 for basic interest scales.

The one exception to this pattern of results was that for turnover intentions, corrected validities for job and vocation focused scales ($-.20$) were only slightly larger than corrected validities for construct focused scales ($-.19$).⁴ However, these results should be interpreted cautiously, given that they are based on only four, multisample military studies as well as on only one job and vocation focused scale and one construct focused scale. Overall, these results provide support for Hypothesis 2 and suggest that job and vocation focused interest scales tend to yield larger validity

estimates than the single, most theoretically relevant construct and basic interest scale for a given job.

Approach to linking multiple interests to criteria. Hypothesis 3 predicted that larger validity estimates would be obtained using regression-weighted composites than using congruence indices. The relevant validities are presented in Table 2. Because the sampling distribution for *R* is unknown (e.g., due to varying numbers of predictors across studies), we do not present the percent of variance explained by artifacts, confidence intervals, or credibility intervals for the composite-level validities. For the criterion of job performance, the observed and corrected *R*s were .28 and .37, respectively, and the shrunken *R*s were .14 and .18. In contrast, the observed and corrected validities for the congruence indices were .12 and .15. For turnover, the observed and corrected *R*s were $-.25$ and $-.33$, and the shrunken *R*s were $-.21$ and $-.29$. However, excluding an influential case reduced the observed and corrected *R*s to $-.22$ and $-.23$ and the shrunken *R*s to $-.13$ and $-.15$. In comparison, the observed and corrected validities for the congruence indices were $-.14$ and $-.15$.

These results provide some support for Hypothesis 3 and suggest that using regression-weighted composites to combine information from multiple construct focused or basic interest scales may yield somewhat higher levels of prediction than using congruence indices. However, the magnitude of the differences depends on whether one uses the unshrunk or shrunken *R*s as the basis for comparison. Also, relatively few studies that met our inclusion criteria used congruence indices for prediction, and thus these validity estimates are based on small subsets of primary studies (*k* = 8 and 4 for job performance and turnover, respectively).⁵ Perhaps of greater interest is the pattern of results that

⁴ We reported moderators within each criterion whenever there were at least two studies for each level of the moderator. In this instance, none of the studies that examined turnover intentions as a criterion used a basic interest scale, and thus we could compare only job and vocation focused scales to construct focused scales. Please see the Table 5 notes for subsequent instances in which we could not examine all levels of a particular moderator.

⁵ Although the relatively larger number of regression-based validity estimates may give the impression that this is a common practice in the literature, it is not. For the most part, we estimated these validities using predictor and criterion correlations reported in articles or obtained upon request from the authors.

Table 4

Descriptive Statistics, Weighted Least Squares (WLS) Regression Results, and Predicted Validities for RIASEC Interests Based on Job Relevance

Variable	Job relevance scores		WLS regression results		Predicted validities for different levels of job relevance				
	<i>M</i>	<i>SD</i>	β	<i>t</i>	−2 <i>SDs</i>	−1 <i>SD</i>	<i>M</i>	+1 <i>SD</i>	+2 <i>SDs</i>
Realistic	56.12	32.06	.31	1.78*	−.06	−.03	.01	.04	.07
Investigative	38.83	22.62	.44	2.76**	−.12	−.05	.02	.09	.15
Artistic	22.37	18.18	.37	2.20**	−.17	−.08	.01	.10	.19
Social	63.93	24.00	.37	2.23**	−.11	−.04	.03	.10	.17
Enterprising	66.08	16.93	.35	2.00*	−.05	−.01	.02	.05	.09
Conventional	56.10	16.54	.23	1.31	−.07	−.04	−.01	.03	.06

Note. *k* ranged from 31 to 33 across realistic, investigative, artistic, social, enterprising, and conventional (RIASEC) dimensions. Job relevance scores range from 0 to 100, whereby higher scores indicate that the RIASEC dimension has relevance to the work performed in a given job. β = standardized regression coefficient from regression of observed criterion-related validity estimates (i.e., for the RIASEC dimension in relation to job performance) on job relevance scores. Predicted validities reflect the predicted relationship between the RIASEC dimension and job performance for the given level of job relevance. For example, the validity estimate of −.06 for realistic at −2 *SDs* represents the predicted validity for realistic interests when the job relevance of this interest is 2 *SDs* below the mean (i.e., very low job relevance).

* $p < .10$. ** $p < .05$.

emerged from analyses used to test Hypotheses 2 and 3. Specifically, validity estimates consistently were larger when multiple interests were considered, either within a single job and vocation focused scale, or when multiple, individual construct focused or basic interest scales were combined via regression analysis or congruence indices.

Other potential moderators. We also explored the possible influence of five additional moderator variables about which we did not have specific hypotheses (see Table 5). First, results suggested that the nature of the criterion measure moderated validity for both job performance and turnover. Specifically, corrected validities for job performance were larger for objective performance criteria (.27) than for ratings criteria (.13), and corrected validities for turnover were somewhat larger for voluntary turnover (−.22) than for involuntary turnover (−.15) or a combination of types of turnover (−.11).

Findings were somewhat mixed with regard to study sample. For job performance, corrected validity estimates were larger for applicant samples (.22) than for employee samples (.14), although the applicant mean estimate was based on only four studies. In contrast, corrected validities for turnover were slightly smaller for applicant samples (−.13) than for employee samples (−.16). Once again, however, the number of applicant studies was relatively small. Regarding study design, predictive designs yielded somewhat smaller corrected validity estimates than concurrent designs for job performance (.13 vs. .15; the mean validity for predictive designs decreased to .10 when an influential case was excluded), training performance (.22 vs. .30), and turnover intentions (−.18 vs. −.21).

The next potential moderator we explored was civilian versus military settings. For job performance, civilian studies were associated with larger corrected validities than were military studies (.21 vs. .09), although there was wide variance in corrected validities within civilian studies ($SD_p = .15$). In contrast, civilian studies yielded smaller corrected validities than military studies for training performance (.12 vs. .28), and the mean validities were similar for turnover (−.14 vs. −.16, respectively).

Finally, with regard to publication status, published studies were associated with larger corrected validity estimates than unpublished studies for job performance (.26 vs. .09, respectively) and for turnover (−.19 vs. −.13). Conversely, unpublished studies were associated with *larger* corrected validities than published studies for training performance (.28 vs. .00), although only three published studies used performance during training as a criterion. We also note that publication status was highly related to study setting ($r = .78$; see Table 6), such that civilian studies tended to be published and military studies tended to be unpublished. Thus, the pattern of results for these two moderators was very similar.

Multiple regression analysis of moderator variables. Next, we used weighted least squares regression to explore relationships between the moderator variables and estimates of criterion-related validity. We used job performance as the criterion for this analysis given the relatively larger number of samples available for this criterion. Each moderator was binary-coded (i.e., 0 vs. 1) to represent the two levels of each moderator (please see the note to Table 6 for details regarding these codes). We then regressed the observed interests–job performance validity coefficients on all the moderators, weighting each study by the inverse of the sampling error variance. One study (Wiggins & Weslander, 1986) emerged as a very strong influential case (Cook's $D = 1.00$ vs. a mean D of .01 across the remaining primary studies). Thus, we excluded this study from the final analysis.

Table 6 displays the moderator intercorrelations and regression results. Most of the moderators were statistically significant bivariate predictors of validity. As a group, the moderators accounted for 48% of the variance in observed validities (shrunk $R^2 = .43$). Three moderators appeared to be the strongest predictors within the regression model: type of interest scale, publication status, and type of performance measure. Specifically, controlling for the effects of other variables in the model, validity estimates were predicted to be .14 points higher for job and vocation focused interest scales than for basic

Table 5

Meta-Analytic Results for Coded Moderators of Criterion-Related Validity of Vocational Interests

Moderator/criterion/category	<i>k</i> ^a	<i>N</i>	<i>r</i>	ρ	<i>SD</i> _{ρ}	% VE	95% CI	80% CV
Type of interest scale								
Job performance								
Job and vocation focused scales	26	3,268	.20	.23	.12	40.1	.17, .29	.07, .39
Construct focused scales	37	4,445	.07	.10	.14	39.3	.03, .15	-.08, .27
Without influential case	36	4,322	.06	.07	.07	67.8	.02, .11	-.02, .17
Basic interest scales	23	8,018	.08	.11	.07	44.3	.07, .15	.02, .20
Training performance								
Job and vocation focused scales	7	2,134	.26	.29	.05	53.9	.23, .35	.22, .35
Construct focused scales	3	382	.13	.15	.10	49.8	-.01, .31	.02, .28
Basic interest scales	8	1,285	.20	.23	.02	100.0	.17, .29	.21, .25
Turnover intentions								
Job and vocation focused scales	5	1,572	-.19	-.20	.10	25.7	-.10, -.29	-.07, -.32
Construct focused scales	17	2,150	-.19	-.19	.05	77.8	-.15, -.24	-.13, -.25
Turnover								
Job and vocation focused scales	10	8,977	-.16	-.17	.07	18.9	-.12, -.21	-.08, -.25
Construct focused scales	11	3,694	-.11	-.12	.06	42.1	-.07, -.17	-.04, -.20
Basic interest scales	6	889	-.11	-.13	.00	100.0	-.08, -.18	-.13, -.13
Nature of criterion								
Job performance								
Ratings measures	71	13,385	.10	.13	.12	36.4	.09, .16	-.02, .28
Objective measures	10	1,166	.24	.27	.11	44.3	.18, .37	.13, .42
Turnover								
Voluntary	15	8,027	-.20	-.22	.12	11.1	-.16, -.29	-.08, -.37
Involuntary	2	3,926	-.13	-.15	.00	100.0	-.14, -.16	-.15, -.15
Other turnover ^b	15	5,426	-.11	-.11	.03	76.6	-.08, -.14	-.08, -.15
Study sample ^c								
Job performance								
Applicants	4	266	.20	.22	.11	60.0	.05, .39	.08, .37
Employees	74	13,803	.11	.14	.13	32.3	.10, .17	-.02, .30
Turnover								
Applicants	6	4,321	-.13	-.13	.03	53.5	-.09, -.17	-.09, -.17
Employees	24	10,132	-.15	-.16	.07	29.5	-.12, -.19	-.07, -.25
Study design ^d								
Job performance								
Predictive	30	6,450	.10	.13	.13	29.1	.07, .18	-.03, .30
Without influential case ^e	29	5,862	.08	.10	.08	41.4	.05, .15	-.03, .23
Concurrent	44	7,745	.11	.15	.12	35.8	.09, .19	-.01, .30
Training performance								
Predictive	13	1,930	.19	.22	.04	85.4	.17, .27	.17, .27
Concurrent	4	1,605	.27	.30	.08	30.4	.21, .39	.20, .40
Turnover intentions								
Predictive	7	1,485	-.17	-.18	.03	82.6	-.12, -.23	-.13, -.22
Concurrent	15	2,237	-.20	-.21	.09	46.7	-.15, -.27	-.10, -.32
Study setting ^f								
Job performance								
Civilian	41	5,278	.18	.21	.15	31.2	.16, .27	.02, .40
Military	39	9,244	.06	.09	.06	62.0	.05, .12	.01, .17
Training performance								
Civilian	4	511	.11	.12	.09	55.4	-.01, .25	.00, .23
Military	13	3,024	.25	.28	.03	79.7	.24, .32	.24, .32
Turnover								
Civilian	11	3,670	-.13	-.14	.06	42.9	-.09, -.18	-.06, -.22
Military	19	10,783	-.15	-.16	.07	27.8	-.12, -.19	-.07, -.24
Publication status ^g								
Job performance								
Published ^h	31	3,895	.23	.26	.14	34.1	.20, .32	.08, .44
Unpublished	49	10,627	.06	.09	.06	64.6	.05, .11	.01, .17

Table 5 (continued)

Moderator/criterion/category	<i>k</i> ^a	<i>N</i>	<i>r</i>	ρ	<i>SD</i> _{ρ}	% VE	95% CI	80% CV
Training performance								
Published	3	245	.00	.00	.00	100.0	-.12, .12	.00, .00
Unpublished	14	3,290	.24	.28	.03	82.9	.24, .31	.24, .31
Turnover								
Published	12	5,998	-.18	-.19	.06	33.3	-.14, -.23	-.11, -.26
Unpublished	18	8,455	-.12	-.13	.06	38.8	-.09, -.16	-.05, -.20

Note. *k* = number of validity coefficients from independent samples; *r* = sample-size weighted mean observed validity; ρ = validity estimate corrected for measurement error in the criterion (for job performance, training performance, and turnover intentions) or corrected to a .50 base rate (for turnover); *SD* _{ρ} = standard deviation of ρ ; % VE = percentage of variance in ρ accounted for by sampling error and measurement error in the criterion; 95% CI = lower and upper bounds of the 95% confidence interval for ρ ; 80% CV = lower and upper bounds of the 80% credibility value for ρ .

^a *k*s for some moderator analyses are larger or smaller than the overall *k*. For example, some studies reported multiple validity estimates that contribute to more than one moderator category (e.g., both job and vocation focused scales and basic interest scales), and the results from other studies are based on both concurrent and predictive data (and thus were excluded from the study design moderator analysis). ^b This category comprised studies that (a) included both voluntary and involuntary turnover or (b) did not clearly indicate the type of turnover examined. ^c All training performance studies used samples of trainees, and all turnover intentions studies used employees. Thus, we could not examine study sample as a potential moderator of validity for these two criteria. ^d All studies that used turnover as a criterion were predictive designs, and thus we could not examine study design as a potential moderator for this criterion. ^e See Footnote 4 regarding identification of influential cases. ^f All studies that used turnover intentions as a criterion were conducted in a military setting, and thus we could not examine study setting as a potential moderator for this criterion. ^g All studies that used turnover intentions as a criterion were unpublished, and thus we could not examine publication status as a potential moderator for this criterion. ^h Published studies include referred journal articles, whereas unpublished studies include dissertations, theses, technical reports, and book chapters.

interest scales,⁶ .10 higher for published studies than for unpublished studies, and .14 points higher for objective performance measures than for performance ratings.

Follow-Up Analyses

We conducted two sets of analyses to follow-up on issues noted earlier. In the first analysis, we attempted to assess the amount of range restriction in vocational interest scales and how accounting for this restriction affects estimates of criterion-related validity. In the second analysis, we revisited the idea that interests that are theoretically *least relevant* to the job may be important to consider.

Range restriction in interests scales. Recall that we could not correct for range restriction due to the lack of applicant standard deviations in the primary studies. However, to estimate the amount and effects of range restriction within this literature, we compared the standard deviations reported in the primary studies in our database to standard deviations from normative samples reported in the manuals of the interest scale(s) examined in each study. Unfortunately, we could not estimate range restriction for most of the primary studies. For example, many studies did not report sample standard deviations, and normative standard deviations were not available for many interest scales (e.g., all the job focused scales). Ultimately, we were able to estimate range restriction (i.e., calculate *u* values) for 17 of the 74 primary studies. Because many studies used multiple interests scales (e.g., a scale for each RIASEC dimension), we were able to calculate 119 total range restriction values. The mean *u* value across all scales was .93, and the mean value for the theoretically most relevant scales from these studies (*k* = 17) was .88. Thus, the amount of range restriction in study samples (relative to normative samples) appears modest, even for the theoretically most relevant interests scales, which one would expect to be particularly restricted.

We then corrected the overall mean validity estimates for each criterion for indirect range restriction using the steps outlined by Hunter, Schmidt, and Le (2006; see their Table 2). We used the .88 *u* value in these analyses because most of our analyses focused on

the theoretically most relevant scale from each study. For job performance, the overall mean validity corrected for measurement error in the criterion was .14. Further correcting this value for indirect range restriction yielded a validity of .17. For training performance, the overall mean corrected validity was .26, and the range restriction-corrected validity was .30. The mean corrected validity for turnover intentions was -.19. Further correcting this value for range restriction yielded a validity of -.23. Finally, the mean corrected validity for actual turnover was -.15, and correcting this value for range restriction yielded a validity of -.18.

Thus, accounting for range restriction in this way yielded somewhat larger estimates of operational validity. However, we urge caution in interpreting these results given our use of normative standard deviations to estimate range restriction, which may be somewhat larger than job-specific (i.e., applicant) standard deviations (e.g., Ones & Viswesvaran, 2003). The larger norm-based standard deviations, in turn, may lead to slightly inflated estimates of range restriction and corrected validity. At the same time, most of the available range restriction estimates are for construct focused interest scales. There were only a few estimates for vocation focused scales, and there were no estimates for job focused scales. As such, we urge readers not to view the results presented above as generalizing to all types of interest measures.

Examination of theoretically least relevant interests. Recall that for studies that reported multiple interests, we used the validity estimate for the interest that appeared to be the theoretic-

⁶ Although validity estimates for all three types of interest scales were considered in the regression analysis, due to the required coding of this three-category moderator, beta coefficients were provided for only two types of scales: job and vocation focused scales and basic interest scales. However, information concerning construct focused scales is reflected in the intercept for this model, which was estimated to be .11. This value reflects the predicted validity of construct focused scales when the values of the other moderators are zero (i.e., a published study, in a civilian setting, using a concurrent design, with an employee sample, and a ratings measure of job performance).

Table 6

Results of Weighted Least Squares (WLS) Regression of Interest–Job Performance Criterion-Related Validity Estimates on Coded Moderators

Variable	1	2	3	4	5	6	7	8
Correlations among moderators and criterion-related validity								
1. Criterion-related validity								
2. Job and vocation focused scales	.48**							
3. Construct focused scales	−.32**	−.58**						
4. Basic interest scales	−.15	−.41**	−.51**					
5. Publication status	−.47**	−.59**	.24*	.35**				
6. Study setting	−.39**	−.56**	.09	.50**	.77**			
7. Study design	−.16	.05	−.12	.09	.01	.11		
8. Study sample	.17	.34**	−.20	−.14	−.18	−.23*	.28*	
9. Type of performance measure	.45**	.27**	−.14	−.13	−.38**	−.35**	.19	.46**
WLS regression results								
	<i>B</i>	<i>SE</i>	β	<i>t</i>				
Intercept ^a	.11	.04		2.92**				
Job and vocation focused scales	.14	.04	.46	3.60**				
Basic interest scales	.03	.03	.12	1.01				
Publication status	−.10	.05	−.35	−2.09*				
Study setting	.03	.05	.12	0.66				
Study design	−.02	.02	−.07	−0.82				
Study sample	−.10	.09	−.11	−1.06				
Type of performance measure	.14	.05	.32	2.95**				
<i>F</i> (7, 71) = 9.27, ** <i>R</i> = .69 (.66), <i>R</i> ² = .48 (.43)								

Note. $N = 79$ independent samples. Criterion-related validity = observed validity coefficient for interests (scale-level measures) and job performance; B = unstandardized regression coefficient; SE = standard error of B ; β = standardized regression coefficient. Job and vocation focused scales was coded 0 if it was not a job or vocation focused scale and was coded 1 if it was. The same coding scheme was used for construct and basic interest scales. Publication status was coded 0 if the study was published and was coded 1 if the study was unpublished. Study setting was coded 0 if the study was conducted in a civilian setting and was coded 1 if it was conducted in a military setting. Study design was coded 0 if the study was concurrent and was coded 1 if the study was predictive. Study sample was coded 0 if the sample comprised employees and was coded 1 if it comprised applicants. Type of performance measure was coded 0 for performance ratings and was coded 1 for objective performance measures. R and R^2 values in parentheses are adjusted for shrinkage.

^a See Footnote 6 regarding the interpretation of the statistics for the intercept.

* $p < .05$. ** $p < .01$.

cally most relevant to the work performed in the target job. However, it has been suggested that interests that are theoretically *least relevant* to a job also may be related to criteria but in a negative direction (e.g., Johnson & Hogan, 1981). For instance, employees who possess interests their job does not support may be less satisfied, less motivated to perform, and more likely to withdraw than employees whose job support their interests.

In light of this issue, we conducted two exploratory analyses. First, for 36 independent samples for which we could determine both the theoretically most relevant RIASEC interest and the theoretically least relevant RIASEC interest (e.g., based on the authors' hypotheses or on O*NET), we counted the number of instances in which each set of interests was the most strongly related to the criterion (i.e., had the largest validity coefficient). Results revealed that the most relevant interest had the largest validity estimate in eight studies, the least relevant interest had the large validity estimate in 12 studies, and a "nonrelevant" interest (i.e., one that was neither most nor least relevant) had the largest validity in the remaining 16 studies.

We then meta-analyzed the validity estimates for the theoretically least relevant RIASEC interest in each study ($k = 36$; $N = 3,974$). The resulting mean observed and corrected validities were $-.09$ and $-.12$, respectively. Recall that the observed and corrected

validities for the theoretically most relevant RIASEC interest from each study were $.07$ and $.10$ (see Table 5). This suggests that interest constructs that a given job clearly does not support may possess about as much predictive value as interests the job clearly does support (although the magnitude of both sets of validity estimates is quite small).

Discussion

Hundreds of primary studies, and numerous meta-analyses, have examined the criterion-related validity of selection procedures such as cognitive ability tests, personality assessments, and employment interviews. In contrast, the validity of vocational interests has not been widely considered within the selection literature in recent decades. The present study helps to address this gap in the literature by cumulating research on the criterion-related validity of interest measures. We analyzed results from 74 primary studies (and 141 independent samples) that examined relations between interests and employee performance or turnover. We believe the findings of this research provide important insights about the potential value of interests for personnel selection.

Key Findings

The first key finding that emerged from the meta-analysis is that the criterion-related validity of vocational interests appears to be somewhat stronger than what many researchers may have thought. Perhaps the most conservative estimates of validity come from studies that used a single interest scale, and for studies that used multiple interest scales, the validity for the theoretically most relevant scale. The estimated mean validities (corrected for measurement error in the criterion) for these studies are .14 for job performance, .26 for training performance, $-.19$ for turnover intentions, and $-.15$ for actual turnover. Note that these validity estimates are corrected only for measurement error in the criterion and not for measurement error in the predictor (or for predictor range restriction, which we discuss below). We did not correct for unreliability in the interest measures because we wanted to know how observed interest scores relate to criteria that organizations may wish to influence through personnel selection (i.e., organizations make decisions on observed scores, not on corrected scores). Because of this, our results underestimate the construct-level relations between interests and the criteria, as well as support for theories that predict such relationships.

We also found considerable variability in interest validity estimates across studies, particularly for job performance criteria, and several variables appear to moderate validity. First, the validity of construct focused interest scales (i.e., those designed to measure one RIASEC dimension) varied by the construct's relevance to the job. In general, as the job relevance of a particular interest construct increases, its correlation with job performance becomes more positive. Conversely, as the job relevance of an interest construct decreases, its correlation with performance becomes more negative.

Another key moderator of interest–criterion relations appears to be type of interest scale. Scales designed to predict performance in, or withdrawal from, a particular job or vocation (which often capture multiple interest constructs) tend to be more valid than scales that assess a single interest construct (e.g., one of RIASEC dimensions) or a single basic interest (e.g., interest in a particular type of activity or work). The estimated corrected validities for job and vocation focused scales are .23 for job performance, .29 for training performance, $-.20$ for turnover intentions, and $-.19$ for actual turnover. In contrast, corrected validities for construct focused and basic interest scales consistently are smaller, particularly for job performance (.10 for construct focused scales and .11 for basic interest scales). Thus, use of any single RIASEC or basic interest scale, even if it is theoretically relevant to the job, is not likely to yield highly useful levels of validity for criteria that reflect employee performance or withdrawal.

However, results suggest that using multiple construct focused or basic interests scales can enhance validity. Creating regression-weighted composites appears to be the most promising approach for combining scores from multiple individual interest scales. The unshrunk and shrunk corrected validities for the regression composites are .37 and .18 for job performance, .36 and .17 for training performance, $-.32$ and $-.24$ for turnover intentions, and $-.33$ and $-.29$ for actual turnover. Also, recall that the shrunk *R*s may provide conservative estimates of criterion-related validity because they are based on all interest scales examined in each

primary study, including scales expected to have no relationship with the criterion.

In contrast, congruence indices yield somewhat smaller corrected validities than the regression-weighted composites. The validity evidence for congruence indices is very similar to the evidence for single, theoretically most relevant scales for both job performance (corrected validities = .15 vs. .14, respectively) and turnover (both corrected validities are $-.15$). However, use of congruence indices to combine information from multiple RIASEC scales can yield slightly larger validities than use of the theoretically most relevant RIASEC scale only (corrected validities = .15 vs. .10, respectively). Thus, a key finding of this meta-analysis is that multiple interests may be needed to yield useful levels of criterion-related validity for employee performance and withdrawal. This can be accomplished by using a job and vocation focused scale (as such scales tend to tap into multiple interests) or by creating regression-weighted composites from scores of multiple construct focused or basic interest scales.

We also explored several additional variables that also appear to moderate the validity of vocational interests. First, the nature of the criterion can affect validity, such that corrected validities for job performance are larger for objective performance criteria than for ratings criteria (.27 vs. .13), and corrected validities for turnover are larger for voluntary turnover ($-.22$) than for involuntary or a combination of types of turnover ($-.15$ and $-.12$). Second, validity estimates from concurrent designs tend to be slightly larger than validity estimates from predictive designs (e.g., .15 vs. .10 for job performance). Finally, with respect to job performance, corrected validities are larger in civilian settings than in military settings (.21 vs. .09).

Last, to our knowledge, this is the first study to cumulate data on range restriction in scores from vocational interest measures. The mean range restriction values (*u*) were .88 for the theoretically most relevant interests and .93 across all interests examined, relevant and nonrelevant. Correcting the overall validity estimates for indirect range restriction (using the .88 value) yields validities of .17 for job performance, .30 for training performance, $-.23$ for turnover intentions, and $-.18$ for actual turnover. It is interesting that the range restriction values we found are similar to range restriction values in the low .90s for personality traits such as conscientiousness and emotional stability (Schmidt, Shaffer, & Oh, 2008). These values suggest that the variance in interests in study samples (which primarily comprised current employees in a particular job) is only slightly less than the variance in interests in normative samples. Furthermore, because we used normative rather than job-specific (i.e., applicant) standard deviations, these values actually may overestimate the amount range restriction in interest scores within employee samples. We note some possible implications of this finding below.

Implications for Theory

The above findings have several implications for theory. First, the fact that vocational interests can demonstrate relations with a range of criteria provides support for several key theories. The finding that interests can relate to performance is consistent with theories from vocational psychology (e.g., Holland, 1973; Lent et al., 1994), industrial–organizational psychology (e.g., J. P. Campbell et al., 1993), and cognitive psychology (e.g., Ackerman, 1996)

that identify interests as a likely antecedent of performance (e.g., by enhancing motivation and/or knowledge and skill acquisition). The fact that interests appear to have some relationship with turnover and turnover intentions appears consistent with theory on vocational interests (e.g., Holland, 1973) and PE fit (e.g., Dawis & Lofquist, 1984), which propose that people whose interests are congruent with the work they do will be more satisfied and, in turn, be more likely to choose to continue doing that work.

Second, we found that interests are somewhat more strongly related to training performance than to job performance and withdrawal. This finding may provide support for theories of individual differences in job performance (e.g., J. P. Campbell et al., 1993) that suggest that interests are likely to have the strongest and most direct effect on motivation and the acquisition of job-relevant knowledge and skills. That is, new employees whose interests are congruent with the job may be more motivated to perform during training and, in turn, acquire the knowledge and skills taught in training. In contrast, interests may be somewhat more distal to performance on the job and whether employees remain in or leave a job.

The present results also may have implications for ASA theory (Schneider, 1987). A common explanation for why interests are thought to be weakly related to criteria such as performance is that the range of interests in a given sample is likely to be quite restricted due to ASA processes (e.g., Spokane, Meir, & Catalano, 2000). Applicant samples are thought to comprise people who are attracted to jobs and organizations they believe will support their interests, and incumbent samples are thought to comprise people who were selected because decision makers thought their interests matched those of the job/organization, and who have remained with the organization because their interests are being met by the job/organization.

However, the range restriction analyses we presented indicate that variance differences between study samples and normative samples are relatively modest. This suggests that employees might not self-select into (or remain in) jobs based on their interests and/or that organizations might not select or terminate on interests, at least not to the extent ASA theory would seem to predict. For example, people may accept and remain in jobs for reasons other than being highly interested in the work, such as pay, location, and the opportunity to obtain experience and to develop knowledge and skills. Further, some people simply “drift into” jobs rather than actively pursuing specific ones (Furnham, 2001).

Nonetheless, the minimal restriction of range in study samples does not necessarily mean that ASA processes do not operate on vocational interests. For instance, there may be substantial *mean* differences in interests across jobs and organizations, which would be consistent with ASA theory (Schmidt et al., 2008). It also is important to remember that our range restriction analyses focused on individual, construct-focused interest scales. As such, our examination does not speak to range restriction in *profiles* of interests, which ASA and Holland’s (1973) theory suggest may be more homogeneous within occupations relevant to the general population. Thus, additional research is needed before we can conclude draw strong conclusions regarding the influence of ASA processes on within-job variance in interest profiles.

Implications for Research and Practice

This meta-analysis also has implications for researchers and practitioners. For one, the results point to several situations in which interests appear to hold more promise for predicting criteria of interest to selection researchers. Specifically, there are at least four situations in which (a) the mean corrected validity is relatively larger in magnitude (keeping in mind the validities are not corrected for range restriction), (b) the variance in corrected validities is not overly large, and (c) the validity evidence is based on more than just a few studies. First, interests appear to be related to job performance when using a job and vocation focused scale (corrected validity = .23, $k = 26$) or when combining multiple construct focused or basic interest scales via multiple regression (observed and shrunken corrected $R_s = .37$ and $.18$, $k = 60$). Second, interests appear to be related to performance during training. The corrected, scale-level validity estimate is .26 ($k = 17$), and the observed and shrunken composite-level validities are .36 and .17, respectively ($k = 8$). Third, there appears to be a relationship between interests and turnover intentions. Across 17 independent samples, the corrected, scale-level validity estimate is $-.19$, and the observed and shrunken composite-level validities are $-.32$ and $-.24$. Finally, interests seem to be predictive of voluntary turnover (corrected validity = $-.22$, $k = 15$).

Furthermore, there may be additional situations in which interests may (or may not) have value that we could not examine due to the number and types of primary studies currently available (e.g., validity of interests for particular types of jobs). Thus, we encourage researchers and practitioners to carefully consider whether interests may have relevance given the particulars of their situation. For example, it is critical to try to examine the theoretical or conceptual relevance of different interest constructs given the requirements of a given job and the nature of the criteria to be predicted. If interests appear to have some relevance, then researchers and practitioners might consider the findings of the current meta-analysis to identify “best-bets” in terms of the types of settings, interest measures, and criteria for which interests may hold the most promise.

With regard to identifying “best bets” for interest measurement, the present results suggest that researchers and practitioners should consider using multiple interests. One option here would be to use a job-specific interest measure. This would involve developing a new, multiscale interest measure (for a recent example, see Van Iddekinge et al., 2011) or adopting scales from an existing interest measure (e.g., the Strong Interest Inventory) that differentiate employees with respect to target criteria (e.g., job performance). A limitation of this option is the time and resources needed to develop and validate such a scale. A second option would be to use multiple individual scales from an established interest measure, such as a measure of the RIASEC dimensions. This option may require less time and resources but would require decisions concerning which scales to retain and how to combine and/or weight the scales. It also may require longer administration times if applicants need to complete several scales. Further, using a commercially available interest measure could be more costly over time than developing a new measure.

The present findings also underscore the importance of understanding how different interest constructs are reflected in the work performed in a particular job. Unlike predictor constructs (e.g.,

cognitive ability) and methods (e.g., structured interviews) that tend to be positively related to performance in most jobs, the direction and magnitude of validity coefficients for interests can vary considerably depending on how relevant a particular interest is to the job. In fact, the present results suggest that in addition to identifying interests that are theoretically relevant to the job, it also may be useful to identify interests that are theoretically *least* relevant to the job, as such interests may be as or more valid predictors (albeit negative) of criteria. This has implications for typical job analysis methods that focus only on what a job requires or, in this case, what interests a job supports. If researchers and practitioners wish to consider interests for selection, they may need to take additional steps during the job analysis to determine the range of interests prospective applicants may possess and which interests the job does and does not support.

Limitations and Future Research

We conclude by noting some factors that limit the conclusions we can draw from our work, as well as some ideas for how future research might address these limitations. First, some of the present results are based on small subsets of studies. Methodological issues (e.g., reporting of statistically significant results only, use of extreme group designs) prevented us from including more primary studies in the meta-analysis. We hope our documentation of these issues (see Table 1) will be of use to researchers in the design, conduct, and reporting of their work. For example, we encourage researchers to provide details regarding how they measured and scored interest scales and how they linked scores on the measures to criteria. We feel this is particularly important given the range of ways interests can be used for prediction. We also call on researchers to report all validity results, both significant and nonsignificant, including all intercorrelations.

One of the smaller subsets of data included studies that examined the validity of interests in applicant samples. The dearth of applicant studies is particularly concerning if interest measures are susceptible to response distortion. Although the debate concerning the prevalence and effects of applicant faking on noncognitive selection procedures continues (e.g., Morgeson et al., 2007; Ones, Dilchert, Viswesvaran, & Judge, 2007; Tett & Christiansen, 2007), we are not aware of any recent studies that have investigated response distortion on interest measures (for some early work in this area, see Kirchner, 1961; Mayo & Guttman, 1959; Stephenson, 1962). Clearly, an important need for future research is to further examine the criterion-related validity of interest measures in operational selection settings using predictive designs with job applicants.

Second, for studies that reported validities for multiple interest scales, we used the validity estimate for the interest that appeared theoretically most relevant to the target job (i.e., for inclusion in the scale-level analyses). For most studies, identification of the most relevant scale was clear-cut. However, there were some cases in which this was less clear (e.g., when the primary authors did not identify the most relevant scale), and thus we had to rely on external information (e.g., from O*NET) to make this determination. In future studies, it would be helpful for researchers to clearly delineate the extent to which each interest measured is relevant to the criteria of interest.

Third, none of the studies we reviewed reported the statistics needed to estimate range restriction (i.e., unrestricted and restricted standard deviations). Further, very few studies provided details concerning the nature of possible range restriction, such as whether it was direct or indirect. Thus, we had to estimate what the range restriction-corrected validities may be using normative samples as a point of comparison. We encourage future researchers to provide such information to better understand the nature and effects of range restriction for vocational interest scales.

The present results suggest that certain types of interest measures may be useful for predicting performance on the job. However, very little is known about whether interests are more strongly related to some performance criteria than to others. As an example, certain types of interests (e.g., social interests) may relate more strongly to interpersonal oriented discretionary behaviors (e.g., organizational citizenship) than to task behaviors. Further, there is almost no research on relations between interests and counterproductive or adaptive work behaviors. It would be informative to know whether interests relate to either of these important criteria. Likewise, although theory suggests that interest may affect performance through motivation and job knowledge and skills (e.g., Ackerman, 1996; J. P. Campbell et al., 1993), there is limited empirical evidence concerning these potential mediating mechanisms. Future research might attempt to test these proposed linkages.

More work also is needed to better understand the role interests play in larger prediction models of job performance and turnover. For example, interests do not tend to correlate highly with cognitive ability (e.g., Ackerman & Heggestad, 1997), and thus interests may provide incremental validity beyond this established predictor of job and training performance (e.g., Van Iddekinge et al., 2011). Further, given the prominence of personality variables such as the Big Five factors (e.g., Barrick & Mount, 1991), it is important to examine whether interests increment the validity of personality, and vice versa. For example, some researchers have suggested that whereas interests influence the types of vocations and jobs people choose, personality is more likely to influence people's behavior once in that vocation or job (e.g., Barrick & Mount, 2005; Hogan & Blake, 1996; Low & Rounds, 2005; Sullivan & Hansen, 2004). However, the results of this meta-analysis, as well as the various theoretical perspectives we discussed, clearly suggest that interests can affect on-the-job behavior. Thus, a key direction for future research is to directly compare the validity of interests and personality with respect to criteria such as job performance and turnover.

Finally, future research is needed to understand how interests fare with respect to other outcomes important to selection researchers. For example, past research has found evidence of demographic differences on certain vocational interests. Men, for instance, tend to have higher realistic and investigative interests, whereas women tend to have higher social, artistic, and conventional interests (Su, Rounds, & Armstrong, 2009). In contrast, ethnic group differences on interest tend to be small and often favor minority individuals (e.g., Betz & Wolfe, 2005). Future research could examine the implications of such differences (or lack thereof) for adverse impact in a selection context. Furthermore, we are not aware of any research that has investigated how applicants react to vocational interests measures.

Conclusions

This meta-analysis helps to clarify the potential value of using vocational interest scales to predict criteria important to selection researchers. The findings suggest that interests hold some promise for predicting both employee performance and turnover, particularly performance during training. Validity estimates for scales that assess job or vocation specific interests appear to be especially promising. In contrast, scales of individual RIASEC constructs or basic interests tend to be weakly related to criteria, even when they are theoretically relevant to the target job. Finally, validity estimates are largest when multiple interests are used for prediction, either within a single job or vocation focused scale, or by using several RIASEC or basic interest scales and combining their scores via regression analysis. We hope this study stimulates future research that considers whether and how interests may enhance the prediction and understanding of behavior at work.

References

References marked with an asterisk indicate studies included in the meta-analysis.

- *Abrahams, N. M., & Neumann, I. (1971). *The assessment of career motivation among NROTC applicants with the Strong Vocational Interest Blank* (Technical Report No. SRR 72-9). San Diego, CA: Naval Personnel and Training Research Laboratory.
- Abrahams, N. M., Neumann, I., & Dann, J. E. (1969). *Use of the Strong Vocational Interest Blank in identifying Naval Academy early motivational disenrollees* (Technical Report No. SRR 70-4). San Diego, CA: Naval Personnel and Training Research Laboratory.
- *Abrahams, N. M., Neumann, I., & Githens, W. H. (1967). *The Navy Preference Blank as a predictor of reenlistment* (Technical Report No. STB 68-1). San Diego, CA: U.S. Naval Personnel Research Activity.
- Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence*, 22, 227–257.
- Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, 121, 219–245.
- Arvey, R. D., & Dewhirst, H. D. (1979). Relationships between diversity of interests, age, job satisfaction, and performance. *Journal of Occupational Psychology*, 52, 17–23.
- *Bain, V. J. A. (1977). *Role assimilation by new employees: The effects of vocational interest appropriateness on the development of work role ambiguity and concomitant performance and job satisfaction* (Unpublished doctoral dissertation). University of California, Berkeley.
- Barge, B. N., & Hough, L. M. (1988). Utility of interest assessment for predicting job performance. In L. M. Hough (Ed.), *Literature review: Utility of temperament, biodata, and interest assessment for predicting job performance* (ARI Research Note 88-02; pp. 131–188). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- 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. (2005). Yes, personality matters: Moving on to more important matters. *Human Performance*, 18, 359–372.
- 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., & Zimmerman, R. D. (2005). Reducing voluntary, avoidable turnover through selection. *Journal of Applied Psychology*, 90, 159–166.
- Beal, D. J., Corey, D. M., & Dunlap, W. P. (2002). On the bias of Huffcutt and Arthur's (1995) procedure for identifying outliers in the meta-analysis of correlations. *Journal of Applied Psychology*, 87, 583–589.
- Betz, N. E., & Wolfe, J. B. (2005). Measuring confidence for basic domains of vocational activity in high school students. *Journal of Career Assessment*, 13, 251–270.
- *Bieschke, K. J., Herbert, J. T., & Bard, C. (1998). Using a social cognitive model to explain research productivity among rehabilitation counselor education faculty. *Rehabilitation Education*, 12, 1–16.
- *Bills, M. A. (1938). Relation of scores in Strong's interest analysis blanks to success in selling casualty insurance. *Journal of Applied Psychology*, 22, 97–104.
- Bizot, E. B., & Goldman, S. H. (1993). Prediction of satisfactoriness and satisfaction: An 8-year follow-up. *Journal of Vocational Behavior*, 43, 19–29.
- *Bolavovich, D. J. (1948). Interest tests reduce factory turnover. *Personnel Psychology*, 1, 81–92.
- *Bolavovich, D. J., & Goodman, C. H. (1944). A study of the Kuder Preference Record. *Education and Psychology Measurements*, 4, 315–325.
- *Borman, W. C., Rosse, R. L., Abrahams, N. M., & Toquam, J. L. (1979). *Investigating personality and vocational interest constructs and their relationship with navy recruiter performance* (Technical report). Minneapolis, MN: Personnel Decisions Research Institute.
- *Borman, W. C., Rosse, R. L., & Rose, S. R. (1984). *An inventory battery to predict performance in navy officer recruiting: Development and validation* (Technical Report No. 85-13). San Diego, CA: Navy Personnel and Development Center.
- *Brayfield, A. H., & Marsh, M. M. (1957). Aptitudes, interests, and personality characteristics of farmers. *Journal of Applied Psychology*, 41, 98–103.
- Butts, M. M., & Ng, T. W. H. (2009). Chopped liver? OK. Chopped data? Not OK. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical myths and methodological urban legends: Doctrine, verity, and fable in the organizational and social sciences* (pp. 361–386). New York, NY: Routledge.
- Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 35–70). San Francisco, CA: Jossey-Bass.
- Campbell, J. T., Otis, J. L., Liske, R. E., & Prien, E. P. (1962). Assessments of higher-level personnel: 2. Validity of the overall assessment process. *Personnel Psychology*, 15, 63–74.
- *Clark, K. E. (1961). *Vocational interests of non-professional men*. Minneapolis, MN: University of Minnesota Press.
- *Coffey, W. C., Jr. (1949). *Characteristics determining success in a non-selling position in the distributive trades* (Unpublished doctoral dissertation). University of Minnesota, Minneapolis.
- Comrey, A. L., & High, W. S. (1955). Validity of some ability and interest scores. *Journal of Applied Psychology*, 39, 247–248.
- *Cook, K. W. (1996). *The relationship of Holland's concept of congruence to work related outcomes* (Unpublished doctoral dissertation). University of South Florida, Tampa.
- Cortina, J. M. (2003). Apples and oranges (and pears, oh my!): The search for moderators in meta-analysis. *Organizational Research Methods*, 6, 415–439.
- Craig, D. R. (1925). The Preference-Interest Questionnaire in selecting retail saleswomen. *Journal of Personnel Research*, 3, 366–374.
- Cronbach, L. J., & Gleser, G. C. (1953). Assessing similarity between profiles. *Psychological Bulletin*, 50, 456–473.
- Dann, J. E., & Abrahams (1977). *Occupational scales of the Navy Vocational Interests Inventory: III. Relationship to job satisfaction, "A"*

- school grades, and job performance (Technical Report No. 78-3). San Diego, CA: Navy Personnel Research and Development Center.
- *Daughtry, P. N. (1989). *The relationship of selected personality factors to turnover among restaurant managers* (Unpublished master's thesis). University of North Texas, Denton.
- Dawis, R. V., & Lofquist, L. H. (1984). *A psychological theory of work adjustment*. Minneapolis, MN: University of Minnesota Press.
- Day, S. X., & Rounds, J. (1997). A little more than kin, and less than kind: Basic interests in vocational research and career counseling. *Career Development Quarterly*, 45, 207–220.
- *Dicken, C. F., & Black, J. D. (1965). Predictive validity of psychometric evaluations of supervisors. *Journal of Applied Psychology*, 49, 34–47.
- Dik, B. J. (2005). *Moderators of the Holland-type congruence-satisfaction and congruence-performance relations* (Unpublished doctoral dissertation). University of Minnesota, Minneapolis.
- *DiMichael, S. G. (1949). Work satisfaction and work efficiency of vocational counselors as related to measured interests. *Journal of Applied Psychology*, 33, 319–329.
- *DuBois, P. H., & Watson, R. I. (1950). The selection of patrolmen. *Journal of Applied Psychology*, 34, 90–95.
- *Dunnette, M. D., & Aylward, M. S. (1956). Validity information exchange (No. 9-21). *Personnel Psychology*, 9, 245–247.
- Dunnette, M. D., & Kirchner, W. K. (1960). Psychological test differences between industrial salesmen and retail salesmen. *Journal of Applied Psychology*, 44, 121–125.
- Dyer, E. D. (1987). Can university success and first year job performance be predicted from academic achievement, vocational interest, personality, and biographical measures? *Psychological Reports*, 61, 655–671.
- Edwards, J. R. (2002). Alternatives to difference scores: Polynomial regression analysis and response surface methodology. In F. Drasgow & N. Schmitt (Eds.), *Measuring and analyzing behavior in organizations: Advances in measurement and data analysis* (pp. 350–400). San Francisco, CA: Jossey-Bass.
- Edwards, J. R. (2008). Person–environment fit in organizations: An assessment of theoretical progress. *The Academy of Management Annals*, 2, 167–230.
- Edwards, J. R., & Parry, M. E. (1993). On the use of polynomial regression equations as an alternative to difference scores in organizational research. *Academy of Management Journal*, 36, 1577–1613.
- Edwards, J. R., & Shipp, A. J. (2007). The relationship between person–environment fit and outcomes: An integrative theoretical framework. In C. Ostroff & T. A. Judge (Eds.), *Perspectives on organizational fit* (pp. 209–258). San Francisco, CA: Jossey-Bass.
- *Ferguson, L. W. (1958). Life insurance interest, ability, and termination of employment. *Personnel Psychology*, 11, 189–193.
- *Ferrara, T. F. (1975). *The influence of race and sex on the development of biodata and interest tests for predicting supervisory success* (Unpublished doctoral dissertation). The American University, Washington, DC.
- Frank, H., & Wilcox, C. (1978). Development and preliminary cross-validation of a two-step procedure for firefighter selection. *Psychological Reports*, 43, 27–36.
- Fritzsche, B. A., McIntire, S. A., & Yost, A. P. (2002). Holland type as a moderator of personality–performance predictions. *Journal of Vocational Behavior*, 60, 422–436.
- *Fritzsche, B. A., Powell, A. B., & Hoffman, R. (1999). Person–environment congruence as a predictor of customer service performance. *Journal of Vocational Behavior*, 54, 59–70.
- Furnham, A. (2001). Vocational preference and P-O fit: Reflections on Holland's theory of vocational choice. *Applied Psychology: An International Review*, 50, 5–29.
- *Gadel, M. S., & Kriedt, P. H. (1952). Relationships of aptitude, interest, performance, and job satisfaction of IBM operators. *Personnel Psychology*, 5, 207–212.
- Gati, I. (1991). The structure of vocational interests. *Psychological Bulletin*, 109, 309–324.
- *Gellatly, I. R., Paunonen, S. V., Meyer, J. P., Jackson, D. N., & Goffin, R. D. (1991). Personality, vocational interest, and cognitive predictors of managerial job performance and satisfaction. *Personality and Individual Differences*, 12, 221–231.
- *Ghei, S. (1960). Vocational interests, achievements, and satisfaction. *Journal of Counseling Psychology*, 7, 132–136.
- *Ghiselli, E. E. (1942). The use of the Strong Vocational Interest Blank and the Pressey Senior Classification Test in the selection of casualty insurance agents. *Journal of Applied Psychology*, 26, 793–799.
- *Gottfredson, G. D., & Holland, J. L. (1990). A longitudinal test of the influence of congruence: Job satisfaction, competency utilization, and counterproductive behavior. *Journal of Counseling Psychology*, 37, 389–398.
- *Gough, H. G., Bradley, P., & McDonald, J. S. (1991). Performance of residents in anesthesiology as related to measures of personality and interests. *Psychological Reports*, 68, 979–994.
- Griffith, R. W., Hom, P. W., & Gaertner, S. (2000). A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium. *Journal of Management*, 26, 463–488.
- Grimsley, G., & Jarrett, H. F. (1973). The relation of past managerial achievement to test measures obtained in the employment situation: Methodology and results. *Personnel Psychology*, 26, 31–48.
- *Guinn, N., Vitola, B. M., & Leisey, S. A. (1976). *Background and interest measures as predictors of success in undergraduate pilot training* (Technical Report No. AFHRL-76-9). Brooks Air Force Base, TX: Air Force Human Resources Laboratory.
- *Guinn, N., Willbourn, J. M., & Kantor, J. E. (1977). *Preliminary development and validation of a screening technique for entry into the security police career field* (Technical Report No. AFHRL-77-38). Brooks Air Force Base, TX: Air Force Human Resources Laboratory.
- Harmon, L. W., Hansen, J. C., Borgen, F. H., & Hammer, A. L. (1994). *Strong Interest Inventory: Applications and technical guide*. Stanford, CA: Stanford University Press.
- Harrison, D. A., Newman, D. A., & Roth, P. L. (2006). How important are job attitudes? Meta-analytic comparisons of integrative behavioral outcomes and time sequences. *Academy of Management Journal*, 49, 305–325.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- *Hense, R. L. (2000). *The Big Five and contextual performance: Expanding person–environment fit theory* (Unpublished doctoral dissertation). University of South Florida, Tampa.
- Hesketh, B., McLachlan, K., & Gardner, D. (1992). Work adjustment theory: An empirical test using a fuzzy rating scale. *Journal of Vocational Behavior*, 40, 318–337.
- Hogan, R., & Blake, R. J. (1996). Vocational interests: Matching self-concept with the work environment. In K. R. Murphy (Ed.), *Individual differences and behavior in organizations* (pp. 89–144). San Francisco, CA: Jossey-Bass.
- Holland, J. L. (1959). A theory of vocational choice. *Journal of Counseling Psychology*, 6, 35–45.
- Holland, J. L. (1973). *Making vocational choices: A theory of careers*. Englewood Cliffs, NJ: Prentice Hall.
- Holland, J. L. (1985). *Vocational Preference Inventory manual*. Odessa, FL: Psychological Assessment Resources.
- Holland, J. L. (1994). *Self-Directed Search Assessment booklet*. Odessa, FL: Psychological Assessment Resources.
- Holland, J. L. (1997). *Making vocational choices: A theory of vocational*

- personalities and work environments (3rd ed.). Odessa, FL: Psychological Assessment Resources.
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). New York, NY: Wiley.
- Hough, L. A., Barge, B. N., & Kamp, J. D. (2001). Assessment of personality, temperament, vocational interests, and work outcome preferences. In J. P. Campbell & D. J. Knapp (Eds.), *Exploring the limits in personnel selection and classification* (pp. 111–154). Mahwah, NJ: Erlbaum.
- Huffcutt, A. I., & Arthur, W. (1995). Development of a new outlier statistic for meta-analytic data. *Journal of Applied Psychology*, 80, 327–334.
- 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. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage.
- Hunter, J. E., Schmidt, F. L., & Le, H. (2006). Implications for direct and indirect range restriction for meta-analysis methods and findings. *Journal of Applied Psychology*, 91, 594–612.
- Iachan, R. (1984). A measure of agreement for use with the Holland classification system. *Journal of Vocational Behavior*, 24, 133–141.
- *Ingerick, M., Diaz, T., & Putka, D. J. (2009). *Investigations into Army enlisted classification systems: Concurrent validation report* (Technical Report No. 1244). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- *Johnson, J. A., & Hogan, R. (1981). Vocational interests, personality, and effective police performance. *Personnel Psychology*, 34, 49–53.
- Kahn, J. H., & Scott, N. A. (1997). Predictors of research productivity and science-related career goals among counseling psychology graduate students. *The Counseling Psychologist*, 25, 38–67.
- *Keiffer, K. M., Schinka, J. A., & Curtiss, G. (2004). Person–environment congruence and personality domains in the prediction of job performance and work quality. *Journal of Counseling Psychology*, 51, 168–177.
- Kemery, E. R., Dunlap, W. P., & Griffeth, R. W. (1988). Correction for range restrictions in point-biserial correlations. *Journal of Applied Psychology*, 73, 688–691.
- Kirchner, W. J. (1961). “Real-life” faking on the Strong Vocational Interest Blank by sales applicants. *Journal of Applied Psychology*, 45, 273–276.
- *Kittrell, D. L. (1980). *The relationship among the interest score on six occupational themes and job satisfaction and performance of Ohio Cooperative Extension county agents* (Unpublished doctoral dissertation). Ohio State University, Columbus.
- *Knapp, D. J., & Heffner, T. S. (Eds.). (2009). *Validating future force performance measures (Army Class): End of training longitudinal validation* (Technical Report No. FR-09-16). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- *Knapp, D. J., & Tremble, T. R. (Eds.). (2007). *Concurrent validation of experimental Army enlisted personnel selection and classification measures* (Technical Report No. 1205). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- *Knaft, E. B. (1951). Vocational interests and managerial success. *Journal of Applied Psychology*, 35, 160–163.
- *Krebs, P. J., Smither, J. W., & Hurley, R. B. (1991). Relationship of vocational personality and research training environment to the research productivity of counseling psychologists. *Professional Psychology: Research and Practice*, 22, 362–367.
- *Kriedt, P. H., & Gadel, M. S. (1953). Prediction of turnover among clerical workers. *Journal of Applied Psychology*, 37, 338–340.
- Kuder, G. F. (1939). *Kuder Preference Record–Form A*. Chicago, IL: University of Chicago Bookstore.
- Kuder, G. F. (1977). *Activity interests and occupational choice*. Chicago, IL: Science Research Associates.
- *Kunce, J. T. (1967). Vocational interests and accident proneness. *Journal of Applied Psychology*, 51, 223–225.
- *Lau, A. W., & Abrahams, N. M. (1969). *Area scales of the Navy Vocational Interest Inventory as predictors of school performance and rating assignment* (Technical Report No. SRR 70-1). San Diego, CA: U.S. Naval Personnel Research Activity.
- *Lau, A. W., Lacey, L. A., & Abrahams, N. M. (1969). An analysis of the Navy Vocational Interest Inventory as a predictor of career motivation (Technical Report No. SRR 69-27). San Diego, CA: U.S. Naval Personnel Research and Activity.
- Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, 45, 79–122.
- Low, K. S. D., & Rounds, J. (2005). Vocational interests. In J. C. Thomas, D. L. Segal, & M. Hersen (Eds.), *Comprehensive handbook of personality and psychopathology* (Vol. 1, pp. 251–267). Hoboken, NJ: Wiley.
- Low, K. S. D., Yoon, M., Roberts, B. W., & Rounds, J. (2005). The stability of vocational interests from early adolescence to middle adulthood: A quantitative review of longitudinal studies. *Psychological Bulletin*, 131, 713–737.
- *MacKinney, A. C., & Wolins, L. (1960). Validity information exchange (No. 13-01). *Personnel Psychology*, 13, 443–447.
- *Mallinckrodt, B., & Gelso, C. J. (2002). Impact of research training environment and Holland personality type: A 15-year follow-up of research productivity. *Journal of Counseling Psychology*, 49, 60–70.
- *Mass, G. (1979). *Using judgment and personality measures to predict effectiveness in police work: An exploratory validation study* (Unpublished doctoral dissertation). The Ohio State University, Columbus.
- *Mayeske, G. W. (1964). The validity of Kuder Preference Record scores in predicting forester turnover and advancement. *Personnel Psychology*, 17, 207–210.
- Mayo, G. D., & Guttman, I. (1959). Faking in a vocational classification situation. *Journal of Applied Psychology*, 43, 117–121.
- *McCullah, R. D. (1969). *Prediction of academic performance in the U.S. Navy Hospital Corps School vocational training programs* (Unpublished doctoral dissertation). University of Maryland, College Park.
- McDaniel, M. A., & Jones, J. W. (1988). Predicting employee theft: A quantitative review of the validity of a standardized measure of dishonesty. *Journal of Business and Psychology*, 2, 327–343.
- *McHenry, J. J., Hough, L. M., Toquam, J. T., Hanson, M. A., & Ashworth, S. (1990). Project A validity results: The relationship between predictor and criterion domains. *Personnel Psychology*, 43, 335–354.
- McNatt, D. B. (2000). Ancient Pygmalion joins contemporary management: A meta-analysis of the result. *Journal of Applied Psychology*, 85, 314–322.
- *Meir, E. I., Esformes, Y., & Friedland, N. (1994). Congruence and differentiation as predictors of workers’ occupational stability and job performance. *Journal of Career Assessment*, 2, 40–54.
- Meir, E. I., Hadas, C., & Noyfield, M. (1997). Person–environment fit in small army units. *Journal of Career Assessment*, 5, 21–29.
- *Meir, E. I., & Navon, M. (1992). A longitudinal examination of congruence hypotheses. *Journal of Vocational Behavior*, 41, 35–47.
- *Miner, J. B. (1960). The Kuder Preference Record in management appraisal. *Personnel Psychology*, 13, 187–196.
- *Moberg, P. J. (1996). *Social skills, personality, occupational interests, and job performance: A multivariate structural analysis* (Unpublished doctoral dissertation). University of Illinois, Champaign-Urbana.
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K. R., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology*, 60, 683–729.
- Morris, M. A. (2003). *A meta-analytic investigation of vocational interest-based job fit and its relationship to job satisfaction, performance, and*

- turnover (Unpublished doctoral dissertation). University of Houston, TX.
- Mount, M. K., Barrick, M. R., Scullen, S. M., & Rounds, J. (2005). Higher-order dimensions of the Big Five personality traits and the big six interest types. *Personnel Psychology*, 58, 447–478.
- *Muchinsky, P. M., & Hoyt, D. P. (1974). Predicting vocational performance of engineers from selected vocational interests, personality, and scholastic aptitude variables. *Journal of Vocational Behavior*, 5, 115–123.
- *Nash, A. N. (1966). Development of an SVIB key for selecting managers. *Journal of Applied Psychology*, 50, 250–254.
- *Neumann, I., & Abrahams, N. M. (1971). The development of an SVIB scale for predicting disenrollment from the NROTC (Regular) Program (Technical Report No. SRM 71-10). San Diego, CA: Naval Personnel and Training Research Laboratory.
- Noe, R. A., & Colquitt, J. A. (2002). Planning for training impact: Principles of training effectiveness. In K. Kraiger (Ed.), *Creating, implementing, and managing effective training and development* (pp. 53–79). San Francisco, CA: Jossey-Bass.
- Ones, D. S., Dilchert, S., Viswesvaran, C., & Judge, T. A. (2007). In support of personality assessment in organizational settings. *Personnel Psychology*, 60, 995–1027.
- Ones, D. S., & Viswesvaran, C. (2003). Job-specific applicant pools and national norms for personality scales: Implications for range-restriction corrections in validation research. *Journal of Applied Psychology*, 88, 570–577.
- *Oppler, S. H., McCloy, R. A., Peterson, N. G., Russell, T. L., & Campbell, J. P. (2001). The prediction of multiple components of entry-level performance. In J. P. Campbell & D. J. Knapp (Eds.), *Exploring the limits in personnel selection and classification* (pp. 349–388). Mahwah, NJ: Erlbaum.
- Ostroff, C. (1993). Comparing correlations based on individual-level and aggregated data. *Journal of Applied Psychology*, 78, 569–582.
- Ployhart, R. E. (2006). Staffing in the 21st century: New challenges and strategic opportunities. *Journal of Management*, 32, 868–897.
- Ployhart, R. E., Schneider, B., & Schmitt, N. (2006). *Staffing organizations: Contemporary practice and theory* (3rd ed.). Mahwah, NJ: Erlbaum.
- *Reeves, D. J., & Booth, R. F. (1979). Expressed versus inventoried interests as predictors of paramedical effectiveness. *Journal of Vocational Behavior*, 15, 155–163.
- Richards, J. M., Jr. (1993). Career development: A ten-year longitudinal study in population scientists. *Journal of Career Assessment*, 1, 181–192.
- *Rieder, K. L. (1991). *Vocational interests as predictors of effectiveness in the ministry* (Unpublished doctoral dissertation). University of West Virginia, Morgantown.
- *Robbins, J. E., & King, D. C. (1961). Validity information exchange (No. 14-02). *Personnel Psychology*, 14, 217–219.
- Ronan, W. W. (1964). Evaluation of skilled trades performance predictors. *Education and Psychological Measurements*, 24, 601–607.
- *Rosenberg, N., & Izard, C. E. (1954). Vocational interests of Naval Aviation cadets. *Journal of Applied Psychology*, 38, 354–358.
- Roth, P. L., Huffcutt, A. I., & Bobko, P. (2003). Ethnic group differences in measures of job performance: A new meta-analysis. *Journal of Applied Psychology*, 88, 694–706.
- *Ryan, T. A., & Johnson, B. R. (1942). Interest scores in the selection of salesmen and servicemen: Occupational vs. ability-group scoring keys. *Journal of Applied Psychology*, 26, 543–562.
- Sackett, P. R., & Lievens, F. (2008). Personnel selection. *Annual Review of Psychology*, 59, 419–450.
- *Schenkel, K. F. (1953). *Tabulator operators selection, emphasizing relationships among aptitudes, interests, proficiency, job- and vocational-satisfaction* (Unpublished master's thesis). University of Minnesota, Minneapolis.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel selection: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262–274.
- Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). The impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology*, 71, 432–439.
- Schmidt, F. L., Shaffer, J. A., & Oh, I.-S. (2008). Increased accuracy for range restriction corrections: Implications for the role of personality and general mental ability in job and training performance. *Personnel Psychology*, 61, 827–868.
- Schmitt, N., Cortina, J. M., Ingerick, M. J., & Wiechmann, D. (2003). Personnel selection and employee performance. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Handbook of psychology: Industrial and organizational psychology* (Vol. 12, pp. 77–105). Hoboken, NJ: Wiley.
- Schneider, B. (1987). The people make the place. *Personnel Psychology*, 40, 437–453.
- Schneider, B., Smith, D. B., Taylor, S., & Fleenor, J. (1998). Personality and organizations: A test of the homogeneity of personality hypothesis. *Journal of Applied Psychology*, 83, 462–470.
- Shieh, G. (2008). Improved shrinkage estimation of squared multiple correlation coefficient and squared cross-validity coefficient. *Organizational Research Methods*, 11, 387–407.
- *Shultz, I. T., & Barnabas, B. (1945). Testing for leadership in industry. *Transactions of the Kansas Academy of Science*, 48, 160–164.
- *Spaulding, H. C. (1980). *Predicting police officer performance: The development of screening and selection procedures based on criterion-related validity* (Unpublished master's thesis). University of South Florida, Tampa.
- *Spies, C. J. (1965). *Some non-intellectual predictors of classroom success* (Unpublished doctoral dissertation). Washington University, St. Louis, MO.
- Spokane, A. R., Meir, E. I., & Catalano, M. (2000). Person–environment congruence and Holland's theory: A review and reconsideration. *Journal of Vocational Behavior*, 57, 137–187.
- Steel, P. D., & Kammeyer-Mueller, J. D. (2002). Comparing meta-analytic moderator estimation techniques under realistic conditions. *Journal of Applied Psychology*, 87, 96–111.
- Stephenson, R. R. (1962). Faking “chance” on the SVIB. *Journal of Applied Psychology*, 46, 252–256.
- Stone, T. H., & Athelstan, G. T. (1969). The SVIB for women and demographic variables in the prediction of occupational tenure. *Journal of Applied Psychology*, 53, 408–412.
- Stone, V. W. (1961). The Strong Vocational Interest Blank and occupational proficiency. *Journal of Educational Research*, 55, 138–140.
- *Strong, E. K. (1934). Interests and sales ability. *Personnel Journal*, 13, 204–216.
- Strong, E. K., Jr. (1943). *Vocational interests of men and women*. Stanford, CA: Stanford University Press.
- Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological Bulletin*, 135, 859–884.
- Sullivan, B. A., & Hansen, J. C. (2004). Mapping associations between interests and personality: Toward a conceptual understanding of individual differences in vocational behavior. *Journal of Counseling Psychology*, 51, 287–298.
- Super, D. E., & Roper, S. (1941). An objective technique for testing vocational interests. *Journal of Applied Psychology*, 25, 487–498.
- Taylor, P. J., Russ-Eft, D. F., & Chan, D. W. L. (2005). A meta-analytic review of behavior modeling training. *Journal of Applied Psychology*, 90, 692–709.
- Tett, R. P., & Christiansen, N. D. (2007). Personality tests at the cross-

- roads: A response to Morgeson, Campion, Dipboye, Hollenbeck, Murphy, and Schmitt (2007). *Personnel Psychology*, 60, 968–993.
- Tett, R. P., & Meyer, J. P. (1993). Job satisfaction, organizational commitment, turnover intention, and turnover: Path analyses based on meta-analytic findings. *Personnel Psychology*, 46, 259–293.
- Thompson, C. E. (1947). Selecting executives by psychological tests. *Education and Psychology Measurements*, 7, 773–778.
- *Tiffin, J., & Phelan, R. F. (1953). Use of the Kuder Preference Record to predict turnover in an industrial plant. *Personnel Psychology*, 6, 195–204.
- Tinsley, H. E. A. (2000). The congruence myth: An analysis of the efficacy of the person–environment fit model. *Journal of Vocational Behavior*, 56, 147–179.
- *Tziner, A., Meir, E. I., & Segal, H. (2002). Occupational congruence and personal task-related attributes: How do they relate to work performance? *Journal of Career Assessment*, 10, 401–412.
- *Usdin, E., & Shenk, F. (1973). *Validity of background and interest tests for officer stated career intent* (Technical Report No. 72-58). Brooks Air Force Base, TX: Air Force Systems Command.
- Van Iddekinge, C. H., Putka, D. J., & Campbell, J. C. (2011). Reconsidering vocational interests for personnel selection: The validity of an interest-based selection test in relation to job knowledge, job performance, and continuance intentions. *Journal of Applied Psychology*, 96, 13–33.
- Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, 81, 557–574.
- *Walter, T. L. (1979). *Prediction of sales performance and job survival of sales personnel: A regression model* (Unpublished doctoral dissertation). University of Missouri, Columbia.
- *Welsch, L. A. (1967). *The supervisor's employee appraisal heuristic: The contribution of selected measures of employee aptitude, intelligence, and personality* (Unpublished doctoral dissertation). University of Pittsburgh, PA.
- *Weslander, D. L. (1981). *Vocational interest patterns of social studies teachers and rated performance* (Unpublished doctoral dissertation). University of Delaware, Newark.
- Wiggins, J. D., & Moody, A. (1983). Identifying effective counselors through client–supervisor ratings and personality–environmental variables. *Vocational Guidance Quarterly*, 31, 259–269.
- *Wiggins, J. D., & Weslander, D. L. (1986). Effectiveness related to personality and demographic characteristics of secondary school counselors. *Counselor Education and Supervision*, 26, 26–35.
- *Wollack, L., & Kipnis, D. (1960). *Development of a device for selecting recruiters* (Technical Report No. PF-016-05-003-W2). Washington, DC: U.S. Naval Personnel Research Field Activity.
- Wood, J. (2008). Methodology for dealing with duplicate effects in a meta-analysis. *Organizational Research Methods*, 11, 79–95.
- Zimmerman, R. D. (2008). Understanding the impact of personality traits on individuals' turnover decisions: A meta-analytic path model. *Personnel Psychology*, 61, 309–348.

Appendix
Main Codes and Input Values for Primary Studies in the Meta-Analysis

Study	Type of Publication	Setting	Sample	Design	Type of interest scale	Most relevant scale ^a	Criterion	N	r _{yy} /base rate	r	R
Abrahams & Neumann (1971)	Technical report	Military	Employees	Concurrent	Job focused		Turnover (voluntary)	603	.34	.29	n/a
Abrahams et al. (1967)	Technical report	Military	Employees	Predictive	Job focused		Turnover (voluntary)	971	.12	.06	n/a
Sample 1								373	.15	-.03	n/a
Sample 2								55	.66	.14	n/a
Bain (1977)	Dissertation	Civilian	Applicants	Predictive	Construct focused—congruence index		Job performance (ratings)				
Bieschke et al. (1998)	Journal article	Civilian	Employees	Concurrent	Basic interests	Interest in research	Job performance (objective)	129	.77	.40	n/a
Bills (1938)	Journal article	Civilian	Employees	Predictive	Vocation focused	Life insurance and real estate	Job performance (objective)	588	.66	.32	n/a
Bolanovich (1948)	Journal article	Civilian	Employees	Predictive	Job focused		Turnover (voluntary)	110	.31	.33	n/a
Bolanovich & Goodman (1944)	Journal article	Civilian	Trainees	Predictive	Construct focused	Scientific	Training performance	66	.81	.10	n/a
Borman et al. (1979)	Technical report	Military	Employees	Concurrent	Job focused		Job performance (ratings)	194	.77	.22	n/a
Borman et al. (1984)	Technical report	Military	Employees	Concurrent	Job focused		Job performance (ratings)	132	.45	.16	n/a
Brayfield & Marsh (1957)	Journal article	Civilian	Employees	Both types of designs	Construct focused	Outdoor and mechanical	Job performance (ratings)	50	.85	.18	n/a
Clark (1961)	Book	Military	Trainees	Concurrent	Vocation focused	Vocation for which participants were being trained	Training performance	547	.80	.30	n/a
Sample 1								547	.80	.24	n/a
Sample 2								461	.80	.31	n/a
Sample 3											
Coffey (1949)	Dissertation	Civilian	Employees	Concurrent	Construct focused	Clerical and computational	Job performance (ratings)	94	.85	.07	.30/.00
Cook (1996)	Dissertation	Civilian	Employees	Concurrent	Construct focused—congruence index		Job performance (ratings)	116	.66	.15	n/a
Daughtry (1989)	Dissertation	Civilian	Applicants	Predictive	Construct focused	Social	Turnover (type not specified)	300	.40	.01	.17/.10
Dicken & Black (1965)	Journal article	Civilian	Employees	Predictive	Vocation focused	Production manager	Job performance (ratings)	31	.70	.09	n/a
DiMichael (1949)	Journal article	Civilian	Employees	Predictive	Construct focused	Social service	Job performance (ratings)	100	.66	-.02	n/a
DuBois & Watson (1950)	Journal article	Civilian	Employees	Predictive	Vocation focused	Police officer	Training performance	129	.80	.02	n/a
							Job performance (ratings)	129	.66	-.01	n/a
Dunnette & Aylward (1956)	Journal article	Civilian	Not specified	Not specified	Vocation focused	Engineer	Job performance (ratings)	33	.66	.03	n/a
Ferguson (1958)	Journal article	Civilian	Applicants	Predictive	Vocation focused	Life insurance	Turnover (type not specified)	520	.33	.09	n/a
Ferrara (1975)	Dissertation	Civilian	Employees	Concurrent	Job focused		Job performance (ratings)	115	.69	.17	n/a
Fritzsche et al. (1999)	Journal article	Civilian	Employees	Concurrent	Vocation focused	Social ^b	Job performance (ratings)	81	.66	.05	.37/.26
					Vocation focused—congruence index			81	.66	.18	n/a

(Appendix continues)

Appendix (continued)

Study	Type of Publication	Setting	Sample	Design	Type of interest scale	Most relevant scale ^a	Criterion	N	$r_{xy}/\text{base rate}$	r	R
Gadel & Kriedt (1952)	Journal article	Civilian	Employees	Concurrent	Job focused	Administrative	Job performance (ratings)	193	.66	.08	n/a
Gellatly et al. (1991)	Journal article	Civilian	Employees	Concurrent	Basic interests		Job performance (ratings)	59	.66	.22	.32/.14
Ghei (1960)	Journal article	Civilian	Employees	Predictive	Job focused		Job performance (ratings)	53	.66	.37	n/a
Sample 1								46	.66	.39	n/a
Sample 2											
Ghiselli (1942)	Journal article	Civilian	Employees	Concurrent	Vocation focused	Life insurance	Job performance	29	.66	.22	n/a
							Ratings				
							Objective	29	.73	-.01	n/a
Gottfredson & Holland (1990)	Journal article	Civilian	Employees	Predictive	Construct focused	Conventional	Turnover (type not specified)	339	.44	.04	n/a
Gough et al. (1991)	Journal article	Civilian	Employees	Predictive	Vocation focused	Medical	Job performance (ratings)	94	.89	.10	n/a
						technologist					
						Investigative		94	.89	.04	n/a
						Medical service		94	.89	.14	n/a
Guinn et al. (1976)	Technical report	Military	Trainees	Predictive	Basic interests	Air force officer	Turnover (both types)	593	.35	.15	.45/.42
Guinn et al. (1977)	Technical report	Military	Trainees	Predictive	Vocation focused		Turnover (type not specified)	4,501	.08	n/a ^c	.28/.27
Hense (2000)	Dissertation	Civilian	Employees	Concurrent	Basic interests		Job performance (ratings)	129	.66	.06	n/a
					Construct focused—congruence index						
Ingerick et al. (2009)	Technical report	Military	Employees	Concurrent	Construct focused						
Sample 1						Realistic	Job performance (ratings)	68	.45	.16	.49/.41
Sample 2							Turnover intentions	76	.95	.15	.25/.00
Sample 3						Realistic	Job performance (ratings)	55	.45	.15	.34/.07
Sample 4							Turnover intentions	65	.96	.40	.56/.50
Sample 5						Realistic	Job performance (ratings)	10	.45	-.10	.83/.09
							Turnover intentions	12	.96	-.25	.52/.00
						Realistic	Job performance (ratings)	42	.45	.17	.43/.22
							Turnover intentions	75	.96	.34	.56/.51
						Social	Job performance (ratings)	40	.45	.08	.40/.09
Johnson & Hogan (1981)	Journal article	Civilian	Trainees/ Employees	Concurrent	Construct focused	Realistic	Turnover intentions	75	.96	.12	.33/.18
Sample 1											
Sample 2							Training performance	50	.80	-.17	.43/.27
Sample 3							Job performance (ratings)	50	.96	.18	.30/.00
Kieffer et al. (2004)	Journal article	Civilian	Employees	Concurrent	Construct focused—congruence index		Job performance (ratings)	38	.90	-.07	.46/.25
					Construct focused		Job performance (ratings)	514	.66	.09	n/a
Kittrell (1980)	Dissertation	Civilian	Employees	Concurrent	Construct focused	Social	Job performance (ratings)	212	.66	.05	.24/.17
Knapp & Hefner (2009)	Technical report	Military	Employees	Predictive	Construct focused	Realistic					
Sample 1							Job performance (ratings)	248	.45	-.06	.16/.04
							Turnover intentions	343	.94	.17	.25/.21
Sample 2							Turnover (both types)	860	.31	.10	.12/.09
Sample 3							Job performance (ratings)	33	.45	.09	.47/.21
							Turnover intentions	31	.94	.45	.71/.63
							Job performance (ratings)	272	.45	.03	.19/.12
							Turnover intentions	292	.94	.06	.14/.00
							Turnover (both types)	345	.22	.04	.09/.00

(Appendix continues)

Appendix (continued)

Study	Type of Publication	Setting	Sample	Design	Type of interest scale	Most relevant scale ^a	Criterion	N	$r_{yy}/\text{base rate}$	r	R
Sample 4							Job performance (ratings)	368	.45	-.10	.17/.11
							Turnover intentions	551	.94	.20	.28/.26
Sample 5							Turnover (both types)	530	.20	.13	.19/.16
							Job performance (ratings)	82	.45	.21	.25/.00
							Turnover intentions	112	.94	.20	.32/.23
Sample 6							Turnover (both types)	145	.27	.08	.17/.00
							Job performance (ratings)	51	.45	-.33	.45/.31
							Turnover intentions	108	.94	.12	.30/.19
Sample 7							Turnover (both types)	109	.33	.27	.27/.14
							Job performance (ratings)	23	.45	-.25	.38/.00
							Turnover intentions	48	.94	.33	.59/.51
							Turnover (both types)	129	.32	.12	.35/.28
Knapp & Tremble (2007)											
Sample 1	Technical report	Military	Employees	Concurrent	Construct focused	Realistic	Job performance (ratings)	169	.45	.15	.17/.00
							Turnover intentions	183	.91	.17	.26/.19
Sample 2							Job performance (ratings)	40	.45	-.08	.34/.00
							Turnover intentions	43	.91	.25	.57/.47
Sample 3							Job performance (ratings)	39	.45	.06	.30/.00
							Turnover intentions	45	.91	.38	.57/.48
Sample 4							Job performance (ratings)	28	.45	.21	.71/.62
							Turnover intentions	43	.91	.10	.53/.41
Sample 5							Job performance (ratings)	39	.45	.14	.45/.24
							Turnover intentions	48	.91	.41	.51/.40
Knauft (1951)	Journal article	Civilian	Applicants	Predictive	Job focused		Job performance (objective)	37	.77	.53	n/a
Krebs et al. (1991)	Journal article	Civilian	Employees	Concurrent	Construct focused	Investigative	Job performance (objective)	260	.77	.15	n/a
Kriedt & Gadel (1953)	Journal article	Civilian	Employees	Predictive	Job focused		Turnover (voluntary)	358	.30	.19	n/a
Kunze (1967)	Journal article	Civilian	Employees	Concurrent	Vocation focused	Aviator and banker	Job performance (objective)	62	.77	.36	n/a
Lau & Abrahams (1969)	Technical report	Military	Trainees	Predictive	Basic interests		Training performance				
Sample 1							Health services	239	.80	.31	.37/.32
Sample 2							Electronics	97	.80	.15	.27/.00
Sample 3							Office work	93	.80	.14	.21/.00
Sample 4							Office work	77	.80	.23	.42/.26
Sample 5							Mechanical	94	.80	.21	.28/.00
Sample 6							Mechanical	73	.80	.14	.35/.00
Sample 7							Mechanical	168	.80	.06	.27/.14
Lau et al. (1969)	Technical report	Military	Employees	Predictive	Basic interests		Turnover (voluntary)				
Sample 1							Health services	223	.09	.05	.12/.00
Sample 2							Electronics	110	.15	.17	.22/.00
Sample 3							Office work	174	.15	.15	.20/.09
Sample 4							Mechanical	109	.17	.12	.23/.00
Sample 5							Mechanical	64	.30	-.04	.35/.00
Sample 6							Mechanical	209	.48	.15	.23/.10

(Appendix continues)

Appendix (continued)

Study	Type of Publication	Setting	Sample	Design	Type of interest scale	Most relevant scale ^a	Criterion	N	$r_{yy}/\text{base rate}$	r	R
MacKinney & Wolins (1960)	Journal article	Civilian	Applicants	Predictive	Construct focused Vocation focused	Mechanical Group III scale	Job performance (ratings)	43	.66	.15	n/a
Mallinckrodt & Gelso (2002)	Journal article	Civilian	Employees	Predictive	Basic interests	Interests in research	Job performance (objective)	43 324	.66 .77	-.09 .29	n/a n/a
Mass (1979)	Dissertation	Civilian	Employees	Concurrent	Construct focused	Investigative	Job performance (ratings)	325	.77	.12	n/a
Mayeske (1964)	Journal article	Civilian	Employees	Predictive	Construct focused	Realistic	Turnover (type not specified)	18	.97	-.05	n/a
McCullah (1969)	Dissertation	Military	Trainees	Predictive	Construct focused Vocation focused	Outdoor Hospital corpsmen	Training performance	125	.30	.19	.32/.16
Sample 1								108	.94	.34	n/a
Sample 2								76	.94	.23	n/a
McHenry et al. (1990)	Journal article	Military	Employees	Concurrent	Basic interests	Combat	Job performance (ratings)	491	.58	.20	.33/.26
Sample 1						Combat		464	.58	.05	.21/.00
Sample 2						Combat		394	.46	.06	.29/.17
Sample 3						Electronics		289	.60	.09	.26/.00
Sample 4						Mechanics		478	.60	.04	.24/.11
Sample 5						Clerical/Admin		427	.45	-.03	.29/.19
Sample 6						Vehicle operation		507	.58	.02	.27/.18
Sample 7						Medical services		392	.60	.09	.29/.17
Sample 8						Law enforcement		597	.63	.01	.20/.06
Sample 9								774	.30	.16	n/a
Meir et al. (1994)	Journal article	Civilian	Applicants	Predictive	Vocation focused— congruence index		Turnover (type not specified)				
Meir & Navon (1992)	Journal article	Civilian	Applicants	Predictive	Construct focused— congruence index						
Sample 1								48	.66	.32	n/a
Sample 2								62	.23	-.10	n/a
								42	.66	.36	n/a
Miner (1960)	Journal article	Civilian	Employees	Concurrent	Basic interests	Supervision	Turnover (type not specified)	57	.26	.23	n/a
						Persuasive	Job performance (ratings)	420	.85	.14	n/a
Moberg (1996)	Dissertation	Civilian	Employees	Concurrent	Construct focused	Enterprising	Job performance (ratings)	251	.66	-.06	.25/.21
Muchinsky & Hoyt (1974)	Journal article	Civilian	Employees	Concurrent	Vocation focused	Group II scale	Job performance (ratings)	89	.66	.34	n/a
Nash (1966)	Journal article	Civilian	Employees	Concurrent	Job focused		Job performance (ratings)	234	.85	.33	n/a
Neumann & Abrahams (1971)	Technical report	Military	Applicants	Predictive	Job focused	Turnover	Turnover				
						Voluntary		2,051	.25	.18	n/a
						Involuntary		1,954	.21	.12	n/a
						Overall		2,608	.41	.14	n/a
Oppler et al. (2001)	Book chapter	Military	Employees	Predictive	Basic interests		Job performance (ratings)				
Sample 1						Combat		521	.56	.07	.25/.14
Sample 2						Combat		557	.56	.10	.24/.13
Sample 3						Combat		119	.56	-.07	.46/.17
Sample 4						Combat		323	.56	.07	.35/.24

(Appendix continues)

Appendix (continued)

Study	Type of Publication	Setting	Sample	Design	Type of interest scale	Most relevant scale ^a	Criterion	N	$r_{yy}/\text{base rate}$	r	R
Sample 5						Electronics		116	.56	.09	.47/.16
Sample 6						Mechanics		242	.56	.13	.30/.00
Sample 7						Clerical/Administration		214	.56	.09	.37/.17
Sample 8						Vehicle operation		223	.56	.06	.33/.06
Sample 9						Medical services		440	.56	.07	.21/.00
Sample 10						Law enforcement		257	.56	.07	.38/.25
Reeves & Booth (1979)	Journal article	Military	Trainees	Predictive	Vocation focused	Hospital corpsmen	Turnover				
							Voluntary	1,911	.15	.32	n/a
							Involuntary	1,972	.18	.14	n/a
							Overall	2,354	.31	.21	n/a
Rieder (1991)								70	.39	.12	.28/.00
Robbins & King (1961)	Dissertation	Civilian	Employees	Concurrent	Construct focused	Social	Job performance (ratings)	538	.30	.10	n/a
Rosenberg & Izard (1954)	Journal article	Military	Trainees	Concurrent	Construct focused	Persuasive	Turnover (type not specified)	274	.50	.34	n/a
Ryan & Johnson (1942)	Journal article	Civilian	Employees		Construct focused	Mechanical	Turnover (voluntary)				
Sample 1				Both types of designs	Job focused		Job performance (ratings)	75	.80	.29	n/a
				Predictive							
Sample 2				Concurrent				336	.80	.16	n/a
Schenkel (1953)	Thesis	Civilian	Employees		Vocation focused	Tabulating machine operators	Job performance (ratings)	167	.90	.18	n/a
Shultz & Barnabas (1945)	Journal article	Civilian	Employees	Both types of designs	Vocation focused	Production manager	Job performance (ratings)	30	.74	.38	.47/.42
Spaulding (1980)	Thesis	Civilian	Trainees	Predictive	Vocation focused	Police officer	Training performance	266	.80	.20	n/a
			Employees				Job performance (ratings)	233	.80	-.04	n/a
							Training performance	266	.80	.20	n/a
							Job performance (ratings)	233	.80	-.05	n/a
Spies (1965)						Realistic	Training performance	444	.65	.22	n/a
Strong (1934)	Dissertation	Military	Trainees	Predictive	Basic interests	Electronics	Job performance (objective)				
Sample 1	Journal article	Civilian	Employees		Vocation focused	Life insurance		102	.78	.10	n/a
Sample 2			Applicants	Not specified				20	.78	.40	n/a
Sample 3			Employees	Predictive				37	.78	.52	n/a
Tiffin & Phelan (1953)	Journal article	Civilian	Employees	Not specified	Job focused		Turnover (voluntary)	487	.19	.19	n/a
Tziner et al. (2002)	Journal article	Military	Employees	Predictive	Vocation focused—congruence index		Job performance (ratings)	254	.45	.08	n/a
Usdin & Shenk (1973)	Technical report	Military	Trainees	Concurrent	Vocation focused	Air force officer	Turnover intentions				
Sample 1								311	.94	.06	n/a
Sample 2								321	.94	.24	n/a
Sample 3								724	.94	.24	n/a
Sample 4								108	.94	.34	n/a
Sample 5								51	.94	-.08	n/a
Walter (1979)	Dissertation	Civilian	Applicants	Predictive	Vocation focused	Group IX	Job performance (objective)	166	.81	.11	n/a

(Appendix continues)

Appendix (continued)

Study	Type of Publication	Setting	Sample	Design	Type of interest scale	Most relevant scale ^a	Criterion	N	$r_{yy}/\text{base rate}$	r	R
Welsch (1967)	Dissertation	Civilian	Employees	Concurrent	Vocation focused ^d		Job performance (ratings)	61	.66	n/a	.10/.00
Sample 1								73	.66		.09/.00
Sample 2								39	.66		.21/.00
Sample 3								51	.66		.32/.22
Sample 4								99	.66		.26/.20
Sample 5								163	.66		.13/.00
Sample 6								37	.66		.33/.17
Sample 7								121	.66		.24/.18
Sample 8								30	.66		.35/.15
Sample 9								26	.66		.32/.00
Sample 10								45	.66		.44/.37
Sample 11								37	.66		.38/.27
Sample 12								57	.84	.26	.49/.39
Weslander (1981)	Dissertation	Civilian	Employees	Concurrent	Construct focused	Social	Job performance (ratings)	123	.66	.64	.90/.90
Wiggins & Weslander (1986)	Journal article	Civilian	Employees	Concurrent	Construct focused	Social	Job performance (ratings)	260	.80	.21	.23/.20
Wollack & Kipnis (1960)	Technical report	Military	Employees	Predictive	Construct focused	Persuasive	Job performance (ratings)				

Note. r_{yy} = values used to correct the validity estimates. Values for job performance, training performance, and turnover intentions reflect reliability estimates, whereas values for turnover reflect the base rate of turnover in each sample. r = observed validity coefficient. R = multiple correlation between interests and the criterion. R s before the backslash are observed, and R s after the backslash are adjusted for shrinkage. "n/a" means it was not possible to estimate an R given the nature of the interest measure or data available. The values in the table served as the basis for most of the analyses. Values used in some of the secondary analyses (e.g., values for less theoretically relevant realistic, investigative, artistic, social, enterprising, and conventional [RIASEC] dimensions) are not shown to conserve space but are available from the first author upon request.

^a No "most relevant" scales are listed for job focused scales or for congruence indices because these predictors comprise items (job focused scales) or scales (congruence indices) linked to various interests. ^b When multiple types of scales are listed, the codes and data for the first scale were used in the majority of analyses. ^c The authors of this study reported a composite-level validity estimate but no scale-level validities. ^d We could not determine the most relevant scale for the samples from this study. As such, only composite-level validities from this study contributed to the meta-analysis.

Received May 24, 2010
Revision received March 17, 2011
Accepted March 31, 2011 ■