Personality and Job Performance: The Big Five Revisited

Gregory M. Hurtz
University at Albany, State University of New York

John J. Donovan
Virginia Polytechnic Institute and State University

Prior meta-analyses investigating the relation between the Big 5 personality dimensions and job performance have all contained a threat to construct validity, in that much of the data included within these analyses was not derived from actual Big 5 measures. In addition, these reviews did not address the relations between the Big 5 and contextual performance. Therefore, the present study sought to provide a meta-analytic estimate of the criterion-related validity of explicit Big 5 measures for predicting job performance and contextual performance. The results for job performance closely paralleled 2 of the previous meta-analyses, whereas analyses with contextual performance showed more complex relations among the Big 5 and performance. A more critical interpretation of the Big 5–performance relationship is presented, and suggestions for future research aimed at enhancing the validity of personality predictors are provided.

During the several decades prior to the 1990s, the use of personality testing in employee selection was generally looked down on by personnel selection specialists. This was primarily due to pessimistic conclusions drawn by researchers such as Guion and Gottier (1965) in their qualitative review of the personality testing literature and by Schmitt, Gooding, Noe, and Kirsch (1984) in their quantitative meta-analysis of various personnel selection techniques. The general conclusion drawn by these researchers was that personality tests did not demonstrate adequate predictive validity to qualify their use in personnel selection. In fact, Schmitt et al. (1984) found that personality tests were among the least valid types of selection tests, with an overall mean sample-size weighted correlation of .21 for predicting job performance, and concluded that "personality tests have low validity" (p. 420).

Over the past several years, however, there has been an increased sense of optimism regarding the utility of personality tests in personnel selection (Behling, 1998; Goldberg, 1993; Hogan, Hogan, & Roberts, 1996; Hogan & Ones, 1997; Mount & Barrick, 1995). In recent years, researchers have suggested that the true predictive validity of personality was obscured in earlier research by the lack of a common personality framework for organizing the traits being used as predictors (Barrick & Mount, 1991; Hough, 1992; Mount & Barrick, 1995; Ones, Mount, Barrick, & Hunter, 1994). With increasing confidence in the robustness of the five-factor model of personality (Digman, 1990; Goldberg, 1993; John,

1990), researchers in the early 1990s began to adopt this Big Five framework for selection research (Barrick & Mount, 1991; Tett, Jackson, & Rothstein, 1991).

Early meta-analytic work by Barrick and Mount (1991) and Tett et al. (1991) provided evidence suggesting that the Big Five might have some degree of utility for selecting employees into a variety of jobs. In both of these reviews, the researchers used studies that provided correlations between any type of personality variable and job performance, categorizing the various personality variables into one of the Big Five dimensions to estimate the strength of these variables' correlation with job performance. Although their results were not altogether consistent (see Ones et al., 1994, and Tett, Jackson, Rothstein, & Reddon, 1994, for a discussion of reasons), the general consensus drawn by researchers and practitioners was that personality does in fact hold some utility as a predictor of job performance. The impact of these studies on raising the status of personality tests in employee selection has been felt throughout the 1990s. Subsequent meta-analyses by Mount and Barrick (1995) and Salgado (1997) have seemed to solidify this newfound status granted to personality, particularly to Conscientiousness. Behling (1998), for example, recently claimed Conscientiousness as one of the most valid predictors of performance for most jobs, second only to general intelligence.

Much of the recent enthusiasm for the Big Five in personnel selection has been based on this body of meta-analytic work, especially the original work of Barrick and Mount (1991). In fact, on the basis of this work, most researchers seem satisfied to conclude that Conscientiousness is a generally valid predictor of job performance and that it represents the primary, if not the sole, personality dimension for use in personnel selection. We feel that it is necessary to revisit and explore the Big Five in this domain for three main reasons.

First, we feel that there are methodological and statistical issues pertaining to past meta-analytic reviews that warrant a critical reanalysis of the research literature that is commonly cited as supporting the criterion-related validity of the Big Five. Second, as several years have passed since the Big Five was adopted as the

dominant personality framework for personnel selection, we feel it would be beneficial to meta-analyze this body of research in which actual measures of the Big Five were correlated with job performance. Third, given recent developments in the research explicating the job performance criterion domain (e.g., Borman & Motowidlo, 1993, 1997; Motowidlo & Van Scotter, 1994; Van Scotter & Motowidlo, 1996), we feel it would be beneficial to meta-analytically explore the relations between the Big Five and these various dimensions of job performance.

Methodological and Statistical Issues in Past Reviews

With respect to prior meta-analytic work examining the utility of the Big Five in personnel selection, we feel that there are two main weaknesses in these reviews that need to be addressed prior to making conclusions about the use of personality for personnel selection. First, it appears that all four major meta-analyses published up to this point (Barrick & Mount, 1991; Mount & Barrick, 1995; Salgado, 1997; Tett et al., 1991) contain a potential threat to construct validity resulting from the methods the researchers used to derive their meta-analytic estimates of criterion-related validity. This threat stems from the fact that these validity coefficients were largely based on studies that used measures that were not designed to explicitly measure the Big Five personality dimensions. Instead, all four of these reviews were based on data from a diverse collection of non-Big Five measures that were classified post hoc into the Big Five categories. Although these were gallant efforts at addressing the relation between the Big Five and job performance given the limited data available in the literature at that time, this post hoc classification procedure has raised some concern in the personnel selection research community over the validity of the results obtained in these past reviews (Hogan et al., 1996; Ones et al., 1994; Salgado, 1997; Tett et al., 1994).

The central issues concerning this classification procedure are the suboptimal levels of interrater agreement in the classification of the various personality scales into the Big Five dimensions and the misclassification of some scales into these dimensions. An inspection of the methods reported by both Barrick and Mount (1991) and Tett et al. (1991) reveals that the level of interrater agreement achieved within each of these reviews is not entirely satisfactory. For example, Barrick and Mount (1991) reported only 83% or better rater agreement on 68% of the classifications, suggesting less than desirable interrater agreement. In light of such difficulties in agreeing on scale classifications, it is not entirely unlikely that errors may have been made in these classifications. As evidence of this problem, Hogan et al. (1996) found that a number of errors had been made in how scales were classified in these early meta-analyses. Additionally, Salgado (1997) indicated that the same scales had been classified into different categories by the different groups of researchers when they conducted their separate meta-analyses. He suggested that this situation arose because there is a degree of ambiguity about how several scales map onto the Big Five, making it difficult to assign them exclusively to one dimension (Salgado, 1997). These facts raise some questions about the accuracy of the classifications and about the degree to which the meta-analytic findings map onto the actual Big Five constructs.

An issue that is related to the classification of scales is the methods used for aggregating validity coefficients within dimen-

sions. When faced with multiple scales categorized into the same dimension from a single study, Barrick and Mount (1991) entered the average correlation across these scales into their meta-analysis. Tett et al. (1991) entered the average absolute value correlation in such instances. As Mount and Barrick (1995) noted, using the average correlation underestimates the validity of the higher order construct to which these scales purportedly belong. Instead, a composite score correlation should be computed to reflect the correlation between the sum of the lower order constructs and the criterion. Mount and Barrick (1995) and Salgado (1997) used this composite-score correlation procedure and demonstrated a resulting increase in the estimated validities of the Big Five. However, the fact still remains that these are only estimates of the validities of actual Big Five measures, because these researchers' studies did not exclusively include correlations from actual Big Five measures. Thus, the degree to which these meta-analyses have provided accurate estimates of the "true" validities of the actual Big Five remains to be seen.

If we accept these previous estimates of the relation between the Big Five and job performance, our second concern then centers around the overwhelmingly positive interpretation of these estimates. As we mentioned previously, Schmitt et al. (1984) suggested that a mean sample-size weighted observed correlation of .21 for personality, averaged across various personality scales without a unified framework, indicated that personality has low validity for predicting job performance. Consistent with this conclusion, the selection community generally looked down on the use of personality as a means of predicting job performance. We find it curious that a number of years later, after the Big Five framework was adopted in subsequent meta-analyses, there have been such positive conclusions concerning the criterion-related validity of Conscientiousness, given that the mean sample-size weighted observed correlations for Conscientiousness were lower than that found by Schmitt et al. (1984; the mean sample-size weighted observed rs for Conscientiousness ranged from .10, Salgado, 1997, to .18, Mount & Barrick, 1995, in these later meta-analyses). In fact, even Barrick and Mount's (1991) estimate of the true correlation for Conscientiousness, after corrections for range restriction and unreliability in both the predictors and criteria, was approximately equal to Schmitt et al.'s uncorrected estimate. Despite these facts, these later reviews met with immediate enthusiasm for the potentially valuable role of Conscientiousness in selection.

In our view, this enthusiasm has resulted from two forces. First, from a theoretical perspective, the Conscientiousness construct does seem to be logically related to job performance. It makes intuitive sense that individuals who have characteristic tendencies to be dependable, careful, thorough, and hardworking should be better performers on the job. It is therefore understandable that so much interest has arisen in this construct as it relates to employee selection. Schmitt et al. (1984), on the other hand, had no specific construct to point to in their analysis, as their validity coefficient was obtained by combining results across a variety of personality variables with no attempt at categorization.

Second, we believe that these validity coefficients for the Big Five have often been interpreted in relative rather than in absolute terms. That is, in these meta-analyses (with the exception of Tett et al., 1991), Conscientiousness has emerged as the most valid of the Big Five, and this has often been interpreted as indicating that Conscientiousness is valid in an absolute sense. On the contrary,

three of the meta-analyses present estimated true correlations for Conscientiousness ranging from .15 to .22 (including statistical corrections for range restriction, predictor unreliability, and criterion unreliability)—correlations that do not fare extremely well when compared to absolute standards that have been used in related research. A meta-analysis by Iaffaldano and Muchinsky (1985), for example, obtained a correlation of .17 between job satisfaction and job performance; this finding has been widely cited as indicating that there is no meaningful relationship between these constructs. Similarly, Cohen (1988) suggested .20 as an approximate standard that should be met for relationships between constructs to be considered meaningful. Furthermore, as we noted previously, Schmitt et al. (1984) concluded that a correlation of .21 was too low to consider personality a useful predictor of job performance. Finally, Mount and Barrick (1995) raised the standards even further by suggesting that validities below .30 are questionable, given the wide range of more valid predictors we have to choose from.

If we were to adopt this .30 standard, only Mount and Barrick (1995) have provided evidence that Conscientiousness may be a valid predictor of job performance in an absolute sense. Whereas Barrick and Mount (1991) and Salgado (1997) found the estimated true correlations between Conscientiousness and job performance to be .22 and .25, respectively, Mount and Barrick (1995) found an estimated overall true validity of .31. It is likely that this higher true validity is due to Mount and Barrick's use of composites score correlations, as we discussed previously (Mount & Barrick, 1995). However, Salgado's lower estimate of .25 was based on the use of composite score correlations as well and was also based on corrections for predictor unreliability that Mount and Barrick did not perform. Thus, in our view, these findings still do not give definitive estimates of the true validities of explicit Big Five measures and do not allow for confident conclusions regarding the validity of Conscientiousness in an absolute rather than a relative sense. At best, they indicate a low to moderate criterion-related validity for Conscientiousness, despite recent enthusiasm that seems to suggest a much stronger role for Conscientiousness in personnel selection (e.g., Behling, 1998).

Developments in the Explication of the Job Performance Criterion Domain

Another potential area in which the current body of meta-analytic work can be improved on is the treatment of the criterion domain. Barrick and Mount (1991) performed a number of moderator analyses for different types of criterion measures, and the most clear finding was that their indicators of Conscientiousness had a somewhat greater impact on subjective ratings than on various types of objective ratings. The results for the other Big Five dimensions were less clear. Salgado (1997) split the criterion domain into subjective ratings, personnel data, and training criteria and again found Conscientiousness to have a somewhat higher impact on subjective ratings than on objective criteria. Mount and Barrick (1995) were more careful to separate out dimensions of performance criteria that were theoretically meaningful with respect to their relation with Conscientiousness, and they did find a pattern of differences showing Conscientiousness to relate to "will do" or motivational factors more strongly than to "can do" or ability factors. However, Mount and Barrick did not present anal-

yses showing the relations between the other Big Five dimensions and those various criteria. Finally, Tett et al. (1991) performed no moderator analyses for criterion types but instead included only correlations computed between personality scales and the criterion dimensions they were hypothesized to predict, and their results were more positive in terms of the impact of Big Five factors other than Conscientiousness on job performance.

The findings of Tett et al. (1991) and Mount and Barrick (1995) do provide some evidence that the link between the Big Five and job performance might be more complex than has recently been suggested, in that their degrees of validity depend on careful selection of theoretically relevant criterion dimensions. Recent work by Motowidlo and Van Scotter (1994; Van Scotter & Motowidlo, 1996) has likewise indicated that the Big Five have differing relations with theoretically linked dimensions of job performance within the task-versus-contextual distinction explicated by Borman and Motowidlo (1993, 1997). This body of work has suggested that personality predictors should have their largest impact on contextual dimensions of job performance. Van Scotter and Motowidlo (1996) showed further that Extraversion and Agreeableness were more strongly related to the interpersonal facilitation component of contextual performance than they were to task performance. Although the magnitudes of these correlations were rather small, this finding does suggest that perhaps the Big Five dimensions other than Conscientiousness take on importance for predicting certain dimensions of job performance—a finding that may have been masked in the earlier meta-analyses. Thus, we feel that the body of meta-analytic evidence relating the Big Five to job performance would benefit from an exploration of their differential relations with task performance and the dimensions of contextual performance.

Summary and Purpose

In summary, we are suggesting that the current body of meta-analytic work investigating the Big Five as predictors of job performance contains some deficiencies that can now be addressed. One major deficiency, in our view, is that all four of the previous meta-analyses (i.e., Barrick & Mount, 1991; Mount & Barrick, 1995; Salgado, 1997; Tett et al., 1991) suffer a potential threat to construct validity in terms of the degree to which their predictors map onto the actual Big Five personality dimensions. This methodological deficiency may have led to inaccurate estimates of the true relation between the Big Five and job performance. The current body of meta-analytic work in this area has provided general hypotheses about the strength of relation between the actual Big Five dimensions and job performance, suggesting that actual Big Five measures of Conscientiousness can be expected to produce criterion-related validities that are low to moderate in magnitude.

In addition to overcoming this deficiency, we believe an exploration of the criterion-related validity of the Big Five for task versus contextual dimensions of job performance would aid in furthering this area of research. Motowidlo and Van Scotter (1994; Van Scotter & Motowidlo, 1996) have begun to present evidence in support of Big Five factors having differential validity with these different components of job performance. Thus, the purpose of the current study is both to meta-analytically summarize the body of research that has developed in recent years where actual

measures of the Big Five were used as predictors of job perfor-
mance and to test the criterion-related validities of the Big Five for
theoretically relevant dimensions of job performance.

Method

Literature Search

We used four separate methods to obtain validity coefficients for the
present review. First, we conducted a computer-based literature search in
PsycLit (1974–1996) and ERIC (1966–1996) using the key words *person-
ality and job performance, personality and training performance, five
factor model*, and *the Big Five*. Second, we conducted a manual search in
the following journals for the period of time from 1985 to 1998: *Academy
of Management Journal, Human Performance, Journal of Applied Psy-
chology, Journal of Management, Journal of Personality and Social Psy-
chology, Organizational Behavior and Human Decision Processes*, and
Personnel Psychology. Third, we hand searched conference programs from
the last four annual conferences (1994–1997) of both the Society for
Industrial and Organizational Psychology (SIOP) and the Academy of
Management for potential articles to be included in the present review.
Finally, we conducted a citation search in which the reference sections
from previously gathered articles were examined to identify any potential
articles that may have been missed by earlier search methods. Using the
selection criteria outlined below, we found 26 studies, yielding 35–45
independent correlations for each of the Big Five dimensions.

Criteria for Inclusion

For a study to be included in the present quantitative review, three
criteria had to be met. First, only studies using actual workers as partici-
pants in the research were included. Second, the study had to include a
personality inventory that was explicitly designed from its inception to
measure the Big Five (i.e., the measure was constructed with the Big Five
as its a priori conceptual basis). Four distinct measures were identified in
the studies collected for the present review: the NEO Personality Inventory
(NEO-PI), including the revised (NEO-PI-R) and five-factor inventory
(NEO-FFI) versions (Costa & McCrae, 1992), Goldberg's Big Five mark-
ers (Goldberg, 1992), the Hogan Personality Inventory (HPI; Hogan &
Hogan, 1995), and the Personal Characteristics Inventory (PCI; Barrick &
Mount, 1993). Finally, the study had to include an explicit measure of job
performance or training performance as the criterion of interest.

Coding of Potential Moderators and Study Characteristics

Consistent with previous meta-analyses (Barrick & Mount, 1991; Sal-
gado, 1997), two study characteristics were coded and treated as potential
moderators of the relations between the Big Five and job performance: type
of worker occupation and type of performance criterion. Each of the
potential moderator variables was independently coded by both Gregory
M. Hurtz and John J. Donovan to ensure accuracy and completeness of
coding. Overall, a high degree of initial agreement (98%) was obtained
between the two independent raters, and divergent ratings were discussed
by the authors until there was an agreement about the proper coding of the
study in question.

Worker occupation. The first characteristic coded for was the occupa-
tion of the workers being examined in the study. A four-category classi-
fication scheme was used to identify the occupation of all research partic-
ipants: sales workers, customer service representatives, managers, and
skilled and semiskilled workers. Approximately 22% (10 of 45) of the
validity coefficients included in the present review came from studies
examining sales jobs, 27% (12 of 45) came from customer service jobs, 9%
(4 of 45) were based on managerial jobs, and 31% (14 of 45) came from
skilled and semiskilled jobs. Approximately 11% (5 of 45) were not

classifiable into one of these categories because of mixed samples or
inadequate information. These studies were therefore excluded from this
set of moderator analyses.

Criterion type. The type of criterion measure used when examining the
predictive validity of the Big Five was also coded as a potential moderator
of the personality–job performance relationship. The criterion domain was
analyzed in two separate ways. First, a two-category classification scheme
was used, with the various criteria categorized as either measures of job
proficiency or measures of training proficiency. Approximately 93% (42
out of 45) of the correlations were based on job proficiency criteria, and 37
of these were based on subjective ratings of job performance. Previous
meta-analyses have analyzed subjective and objective performance mea-
sures separately; in our data set, the objective analysis would have con-
sisted entirely of objective sales data, making it a subset of studies from the
moderator analysis of the sales occupation. We therefore decided to ex-
clude a separate moderator analysis for objective data. The training profi-
ciency category included both ratings of training performance and end-of-
training tests designed to evaluate learning and hands-on demonstration of
skills. As very few training studies were found that used explicit Big Five
measures, only 7% (3 out of 45) of the studies were based on measures of
training proficiency.

Second, we performed a separate analysis by partitioning the criterion
domain into task performance, job dedication, and interpersonal facilitation
(Motowidlo & Van Scotter, 1994; Van Scotter & Motowidlo, 1996). Using
definitions provided by Motowidlo and Van Scotter (1994) and Van
Scotter and Motowidlo (1996), we classified performance criteria such as
technical performance, use of equipment, job knowledge, completion of
specified job duties, and objective performance data as indicators of the
task performance category; ratings of work dedication, effort, persistence,
reliability, self-direction, commitment to objectives, and the like as indi-
cators of the job dedication category; and ratings of interpersonal relations,
cooperation, quality of interactions with others, being courteous, and being
a team player as indicators of interpersonal facilitation. Using these defi-
nitions, we located within our sample 7–12 validity coefficients (across Big
Five dimensions) for the prediction of task performance criteria, 14–17 for
job dedication criteria, and 19–23 for criteria fitting the definition of
interpersonal facilitation.

Computation of Validity Coefficients

Within individual studies, there were instances in which correlation
coefficients from a single sample had to be combined. On the predictor
side, for the HPI, some studies reported separate correlations for the
Ambition and Sociability subscales of the Extraversion dimension and for
the Intellectance and School Success subscales of the Openness/Intellect
dimension, rather than correlations at the dimension level. In these cases,
rather than averaging across the subscales, we computed the composite
score correlation (see Hunter & Schmidt, 1990, pp. 454–463) to estimate
the correlation between the sum of the two lower level subscales and job
performance. When correlations between the lower level subscales were
not provided in a study, we entered the correlations presented in the HPI
manual into the composite score correlation formula. Combining the lower
level scales in this manner is entirely consistent with the fact that these
subscales were derived directly from the dimension-level scales; thus, their
sum directly assesses the dimension-level construct. Therefore, this does
not undermine our purpose of including only explicit Big Five measures.

Similarly on the criterion side, some studies provided correlations be-
tween the Big Five and separate dimensions of job performance without
providing a correlation with the composite criterion score. In these cases,
we again estimated the composite score correlation rather than simply
averaging across performance dimensions. When the correlations between
performance dimensions were not provided, we entered .55 into the com-
posite score formula. We derived this estimate by first computing the
average correlation among dimensions within the studies that did provide

such information and then computing the mean sample-size weighted correlation across these studies. For the separate analyses of task and contextual performance dimensions, we used the same procedure for combining correlations from a single sample that were based on multiple rating scales classified into a common dimension.

When conducting the actual meta-analysis, we used the Hunter-Schmidt validity generalization framework (Hunter & Schmidt, 1990). Using this framework, we obtained the mean sample size-weighted correlations, the estimated true or operational validities corrected for sampling error, range restriction, criterion unreliability, and the estimated true-score correlations with additional corrections for predictor unreliability.

As in the previous meta-analyses we have just reviewed, these corrections had to be made by way of artifact distributions because of a low rate of reporting the statistics that are necessary for applying corrections to the individual coefficients. Two artifact distributions were created for the criterion reliabilities. For analyses in which only subjective ratings of performance were involved, we created a distribution by augmenting the few interrater reliability coefficients obtained from our sample of studies with those presented in Rothstein (1990). This distribution had a mean criterion reliability of .53 (SD = .15). For those analyses in which a combination of objective and subjective criteria were used, we added to the previous distribution the reliabilities presented in our sample for objective criteria and those presented by Hunter, Schmidt, and Judeisch (1990). Adding these reliabilities created a distribution with a mean of .59 (SD = .19). Although this combined distribution is weighted rather heavily with subjective ratings, this is entirely consistent with the fact that approximately 90% of the criteria in our sample of studies were subjective in nature.

For corrections for predictor unreliability, we created separate artifact distributions for each of the Big Five dimensions by augmenting the reliability estimates provided in our sample of studies with those from the inventory manuals. This provided distributions with mean predictor reliabilities ranging from .76 (SD = .08; Agreeableness) to .86 (SD = .04; Emotional Stability). For range restriction corrections, we found very few unrestricted standard deviations reported in the studies for computing the u values. Thus, we used two strategies for obtaining unrestricted standard deviations. First, we attempted to contact the authors of the inventories to obtain standard deviations from unrestricted samples of applicants. Second, following Salgado's (1997) strategy, we used standard deviations provided in the inventory manuals as the unrestricted values. As we did not have enough information to create reliable separate distributions for each of the five dimensions, we created a single artifact distribution for use in all our analyses. This distribution of u values had a mean of .92 (SD = .27). Overall, our artifact distributions were very similar to those used in the previous meta-analyses. Corrections based on these distributions were conducted interactively using software described by Hunter and Schmidt (1990), on the basis of the recommendations of Law, Schmidt, and Hunter (1994).

Results

Overall Validity Coefficients

Table 1 presents the results of the omnibus meta-analysis across occupations and performance criteria. These analyses were based on a range of 35–45 correlations and 5,525–8,083 job applicants and incumbents. The mean sample-size weighted correlations ($\bar{r}$) ranged from .04 to .14 across dimensions and are substantially lower than the mean correlation of .21 found by Schmitt et al. (1984) and very similar to those found by Barrick and Mount (1991; ranging .03–.13) and Salgado (1997; ranging .01–.10). The estimated true validities (ρ_v) for explicit measures of the Big Five ranged from .06 to .20, and the estimated true-score correlations (ρ_c) ranged from .07 to .22. Consistent with Barrick and Mount (1991) and Salgado (1997), the highest validity of the Big Five dimensions was that for Conscientiousness ($\rho_v = .20$), which demonstrated a low to moderate level of validity. The 90% credibility interval for this dimension did not include zero, suggesting the absence of moderators in this estimate of the true validity (Hunter & Schmidt, 1990; Whitener, 1990). Emotional Stability also had a credibility interval that was greater than zero, although its estimated true validity was substantially lower ($\rho_v = .13$).

Validity Coefficients by Occupation

Table 2 presents the results of the moderator analysis for the occupational categories. Despite the lack of moderators indicated by the credibility intervals for Conscientiousness and Emotional Stability in the omnibus analysis, we carried out all moderator analyses for each of the Big Five for the sake of comparison. For all four of the occupational categories, Conscientiousness exhibited the highest estimated true validity. It is interesting to note that despite the indication of no moderators for Conscientiousness, the estimated true validity for this dimension ranged from .15 to .26 across occupations. Its highest validities were for sales ($\rho_v = .26$) and customer service ($\rho_v = .25$) jobs. The magnitudes of these validities are moderate, and those for the remaining Big Five dimensions remained low across all occupations.

It is noteworthy, however, that some of the low validities for the other Big Five dimensions appear to be rather stable, in that their credibility intervals fall above zero. For sales jobs, Emotional Stability ($\rho_v = .13$) and Extraversion ($\rho_v = .15$) appear to have

Table 1
Overall Validity Coefficients by Personality Dimension

Big Five dimension	k	N	$\bar{r}$	$S_{\bar{r}}^2$	S_e^2	S_{meas}^2	S_{res}^2	% VE	ρ_c	ρ_v	$SD_{\rho v}$	90% CV
Conscientiousness	45	8,083	.14	.0161	.0054	.0016	.0091	44	.22	.20	.14	.03
Emotional Stability	37	5,671	.09	.0084	.0065	.0007	.0013	85	.14	.13	.05	.06
Agreeableness	40	6,447	.07	.0108	.0062	.0005	.0041	62	.13	.11	.09	−.01
Extraversion	39	6,453	.06	.0111	.0060	.0004	.0047	57	.10	.09	.10	−.04
Openness to Experience	35	5,525	.04	.0093	.0064	.0002	.0028	70	.07	.06	.08	−.04

Note. k = number of validity coefficients; N = total sample size; $\bar{r}$ = sample-size weighted mean observed validity; $S_{\bar{r}}^2$ = total observed variance in $\bar{r}$; S_e^2 = variance due to sampling error; S_{meas}^2 = variance due to measurement artifacts; S_{res}^2 = residual variance; % VE = percentage of variance accounted for by sampling error and measurement artifacts; ρ_c = true-score correlation; ρ_v = true (operational) validity; $SD_{\rho v}$ = standard deviation of true validity; CV = credibility value (lower bound of credibility interval for ρ_v).

Table 2

Validity Coefficients for Personality Dimensions by Occupational Category

Big Five dimension	k	N	$\bar{r}$	S_r^2	S_e^2	S_{meas}^2	S_{res}^2	% VE	ρ_c	ρ_v	SD_{ρ_v}	90% CV
					Sales							
Conscientiousness	10	1,369	.18	.0117	.0069	.0026	.0021	82	.29	.26	.07	.17
Emotional Stability	7	799	.09	.0082	.0087	.0007	.0000	115	.15	.13	.00	.13
Agreeableness	8	959	.03	.0098	.0084	.0001	.0013	87	.06	.05	.05	−.02
Extraversion	8	1,044	.10	.0117	.0076	.0009	.0033	72	.16	.15	.08	.04
Openness to Experience	6	732	.03	.0150	.0083	.0001	.0067	55	.04	.04	.12	−.12
					Customer service							
Conscientiousness	12	1,849	.17	.0121	.0062	.0023	.0036	70	.27	.25	.09	.13
Emotional Stability	10	1,614	.08	.0052	.0062	.0006	.0000	129	.13	.12	.00	.12
Agreeableness	11	1,719	.11	.0038	.0063	.0011	.0000	193	.19	.17	.00	.17
Extraversion	10	1,640	.07	.0117	.0061	.0004	.0052	56	.11	.11	.11	−.03
Openness to Experience	9	1,535	.10	.0043	.0058	.0010	.0000	158	.17	.15	.00	.15
					Managers							
Conscientiousness	4	495	.11	.0451	.0079	.0011	.0361	20	.19	.17	.28	−.19
Emotional Stability	4	495	.08	.0088	.0080	.0006	.0002	98	.13	.12	.02	.10
Agreeableness	4	495	−.03	.0040	.0081	.0001	.0000	205	−.04	−.04	.00	−.04
Extraversion	4	495	.08	.0045	.0080	.0006	.0000	192	.13	.12	.00	.12
Openness to Experience	4	495	−.02	.0111	.0081	.0000	.0029	74	−.03	−.03	.08	−.13
					Skilled and semiskilled							
Conscientiousness	14	3,481	.10	.0147	.0040	.0009	.0098	33	.17	.15	.15	−.03
Emotional Stability	11	1,874	.06	.0110	.0059	.0003	.0048	56	.09	.08	.10	−.05
Agreeableness	12	2,385	.06	.0103	.0050	.0004	.0049	52	.11	.10	.10	−.04
Extraversion	12	2,385	.00	.0080	.0051	.0000	.0029	63	.01	.01	.08	−.10
Openness to Experience	11	1,874	−.01	.0062	.0059	.0000	.0003	95	−.02	−.02	.03	−.05

Note. k = number of validity coefficients; N = total sample size; $\bar{r}$ = sample-size weighted mean observed validity; S_r^2 = total observed variance in $\bar{r}$; S_e^2 = variance due to sampling error; S_{meas}^2 = variance due to measurement artifacts; S_{res}^2 = residual variance; % VE = percentage of variance accounted for by sampling error and measurement artifacts; ρ_c = true-score correlation; ρ_v = true (operational) validity; SD_{ρ_v} = standard deviation of true validity; CV = credibility value (lower bound of credibility interval for ρ_v).

low but stable true validities. This same general pattern emerged for managerial jobs, although the small number of studies ($k = 4$) located for estimating this true validity may render this finding tenuous. Customer service jobs appear more complex in that Emotional Stability ($\rho_v = .12$), Agreeableness ($\rho_v = .17$), and Openness to Experience ($\rho_v = .15$) exhibited rather low but stable true validities. This may indicate a somewhat more complex pattern of relationships between personality and performance in jobs that involve interpersonal interactions than is captured solely by assessing Conscientiousness. In contrast, the true validity estimates for skilled and semiskilled jobs, which may often involve a smaller interpersonal component of performance, tended to be rather small across all of the Big Five, and these validities appear rather unstable in light of their credibility intervals.

Validity Coefficients by Criterion Type

Table 3 presents the results of the moderator analysis for the separate predictions of job proficiency and training proficiency. For job proficiency, virtually the same pattern and magnitude of validities emerged as was found in the omnibus analysis, which is not surprising given the fact that over 90% of the individual correlations across dimensions involved job proficiency criteria. The small number of correlations summarized for training proficiency renders interpretation of the true validity estimates tenuous,

although Extraversion ($\rho_v = .17$) and Agreeableness ($\rho_v = .18$) had the highest validities.

Table 4 shows the separate analyses of the Big Five as predictors of task performance, job dedication, and interpersonal facilitation. Recent research and theory explicating the dimensionality of the job performance domain has suggested that personality should predict the contextual performance dimensions of job dedication and interpersonal facilitation more strongly than task performance does (Borman & Motowidlo, 1993, 1997; Motowidlo & Van Scotter, 1994; Van Scotter & Motowidlo, 1996). Our analyses show that Conscientiousness predicted all three criteria with approximately the same level of true validity ($\rho_v = .15-.18$), although the credibility intervals indicate that this true validity was only stable for the interpersonal facilitation criterion. Emotional Stability appeared to have a low but very stable true validity across these three criteria ($\rho_v = .13-.16$). For the interpersonal facilitation criterion, Agreeableness ($\rho_v = .17$) rivaled both Conscientiousness ($\rho_v = .16$) and Emotional Stability ($\rho_v = .16$) in its estimated true validity. This supports Van Scotter and Motowidlo's finding that although Agreeableness does not influence task performance, it does appear to influence ratings of interpersonal facilitation. It should be noted, however, that none of these analyses for the task and contextual performance criteria revealed stronger true validities than did the overall performance analyses.

Table 3
Validity Coefficients for Personality Dimensions by Criterion Type

Big Five dimension	k	N	$\bar{r}$	$S_{\bar{r}}^2$	S_e^2	S_{meas}^2	S_{res}^2	% VE	ρ_c	ρ_v	$SD_{\rho v}$	90% CV
					Job performance							
Conscientiousness	42	7,342	.15	.0148	.0055	.0019	.0074	50	.24	.22	.13	.06
Emotional Stability	35	5,027	.09	.0089	.0069	.0007	.0013	85	.15	.14	.05	.07
Agreeableness	38	5,803	.07	.0111	.0065	.0004	.0042	62	.12	.10	.10	‑.02
Extraversion	37	5,809	.06	.0118	.0064	.0003	.0051	57	.09	.09	.11	‑.05
Openness to Experience	33	4,881	.03	.0097	.0068	.0001	.0028	71	.06	.05	.08	‑.05
					Training performance							
Conscientiousness	3	741	.02	.0145	.0041	.0000	.0104	28	.03	.03	.15	‑.16
Emotional Stability	2	644	.06	.0030	.0031	.0003	.0000	111	.09	.08	.00	.08
Agreeableness	2	644	.12	.0049	.0030	.0013	.0006	88	.21	.18	.04	.13
Extraversion	2	644	.12	.0020	.0030	.0012	.0000	207	.19	.17	.00	.17
Openness to Experience	2	644	.08	.0042	.0031	.0007	.0005	88	.14	.13	.03	.08

Note. k = number of validity coefficients; N = total sample size; $\bar{r}$ = sample-size weighted mean observed validity; $S_{\bar{r}}^2$ = total observed variance in $\bar{r}$; S_e^2 = variance due to sampling error; S_{meas}^2 = variance due to measurement artifacts; S_{res}^2 = residual variance; % VE = percentage of variance accounted for by sampling error and measurement artifacts; ρ_c = true-score correlation; ρ_v = true (operational) validity; $SD_{\rho v}$ = standard deviation of true validity; CV = credibility value (lower bound of credibility interval for ρ_v).

Discussion

The main purpose of this study was to provide a confirmatory meta-analysis of the relation between the Big Five and job performance by including only scales that were explicitly designed to measure the Big Five personality dimensions. Our overall results were highly consistent with the original work of Barrick and Mount (1991), in that Conscientiousness was again found to have the highest validity of the Big Five dimensions for overall job performance. Furthermore, our estimated true-score correlation of .22 (and true validity of .20) was virtually identical in magnitude to Barrick and Mount's estimate. This finding alleviates concern that Barrick and Mount's heavily cited results underestimated the overall true validity of the Conscientiousness dimension as a result of their categorization procedures. On the other hand, our findings indicate that at least for single-scale, global Big Five measures, the

Table 4
Validity Coefficients for Personality Dimensions by Criterion Dimension

Big Five dimension	k	N	$\bar{r}$	$S_{\bar{r}}^2$	S_e^2	S_{meas}^2	S_{res}^2	% VE	ρ_c	ρ_v	$SD_{\rho v}$	90% CV
					Task performance							
Conscientiousness	12	2,197	.10	.0138	.0054	.0008	.0076	45	.16	.15	.13	‑.02
Emotional Stability	8	1,243	.09	.0015	.0064	.0007	.0000	463	.14	.13	.00	.13
Agreeableness	9	1,754	.05	.0090	.0051	.0002	.0037	59	.08	.07	.09	‑.05
Extraversion	9	1,839	.04	.0052	.0049	.0002	.0001	98	.07	.06	.02	.04
Openness to Experience	7	1,176	‑.01	.0237	.0060	.0000	.0177	25	‑.01	‑.01	.20	‑.26
					Job dedication							
Conscientiousness	17	3,197	.12	.0203	.0052	.0013	.0139	32	.20	.18	.17	‑.04
Emotional Stability	15	2,581	.09	.0059	.0058	.0007	.0000	109	.14	.13	.00	.13
Agreeableness	17	3,197	.06	.0096	.0053	.0003	.0040	59	.10	.08	.09	‑.03
Extraversion	16	3,130	.03	.0111	.0051	.0001	.0059	47	.05	.05	.11	‑.10
Openness to Experience	14	2,514	.01	.0108	.0056	.0000	.0052	52	.01	.01	.11	‑.13
					Interpersonal facilitation							
Conscientiousness	23	4,301	.11	.0083	.0053	.0010	.0020	76	.18	.16	.07	.07
Emotional Stability	21	3,685	.10	.0046	.0056	.0010	.0000	142	.17	.16	.00	.16
Agreeableness	23	4,301	.11	.0117	.0052	.0012	.0053	55	.20	.17	.11	.03
Extraversion	21	4,155	.06	.0105	.0050	.0004	.0051	52	.11	.10	.11	‑.04
Openness to Experience	19	3,539	.03	.0075	.0054	.0001	.0020	73	.05	.05	.07	‑.04

Note. k = number of validity coefficients; N = total sample size; $\bar{r}$ = sample-size weighted mean observed validity; $S_{\bar{r}}^2$ = total observed variance in $\bar{r}$; S_e^2 = variance due to sampling error; S_{meas}^2 = variance due to measurement artifacts; S_{res}^2 = residual variance; % VE = percentage of variance accounted for by sampling error and measurement artifacts; ρ_c = true-score correlation; ρ_v = true (operational) validity; $SD_{\rho v}$ = standard deviation of true validity; CV = credibility value (lower bound of credibility interval for ρ_v).

validity estimates for Conscientiousness provided by Mount and Barrick (1995) and Salgado (1997) appear to be overestimates. We offer from our results an estimated true criterion-related validity of .20 for actual Big Five measures of Conscientiousness.

It is also noteworthy that Emotional Stability shows rather consistent (although low) levels of criterion-related validity. In addition, the separate analyses for the different occupational categories provide a more complex picture of the validities of the Big Five than do prior reviews, in that the dimensions beyond Conscientiousness begin to show low but rather stable validities for certain occupations. In particular, for jobs involving customer service, Agreeableness, Openness to Experience, and Emotional Stability had low levels of validity (ρ_vs ranging .12–.17) but zero residual variance in the population estimate after the effects of sampling error and measurement artifacts were removed. Similarly for sales and perhaps for managerial jobs, Emotional Stability and Extraversion had rather low but stable validities. A common theme running through customer service, sales, and managerial jobs that differentiates them from skilled and semiskilled jobs is the interpersonal component of performing these jobs; this probably accounts for the more stable validities of these personality dimensions for these types of jobs.

The results of partitioning the criterion domain into task performance, job dedication, and interpersonal facilitation shed further light on these issues. Contrary to our expectations, our analyses showed that partitioning the criterion domain in this manner did not bring about stronger criterion-related validities in comparison with analyzing a general job proficiency category. Conscientiousness predicted all three performance dimensions equally well (ρ_v = .15–.18), and the same was found for Emotional Stability (ρ_v = .13–.16). However, Agreeableness did emerge as a potentially valid predictor, predicting interpersonal facilitation as strongly as did Conscientiousness and Emotional Stability. When one considers the validities and credibility intervals together, Conscientiousness, Emotional Stability, and Agreeableness appear to have a rather stable impact on the interpersonal facilitation criteria, suggesting that perhaps this performance dimension is influenced in a consistent manner by certain personality traits. This is consistent with the previous suggestion that personality may have a more stable impact on jobs that are more interpersonal in nature (e.g., customer service, sales, management). The same statement may be true of job performance dimensions that are more interpersonal in nature.

Although these findings shed some light on the potential impact of personality variables on dimensions of job performance, we now return to the issue of the absolute magnitudes of the estimated true validity coefficients for these variables and the implications of these validities for the utility of the Big Five for personnel selection. In general, our analyses suggest that the validities of the Big Five, including Conscientiousness, tend to be low to moderate in magnitude. One of the major implications of this meta-analysis, then, is that stating that Conscientiousness is a valid predictor of job performance paints an inaccurate image of the true validity of these global Big Five measures in an absolute sense. We suggest that the estimated true validity of .20 for the global Conscientiousness dimension is not as impressive as one would expect, given the recent enthusiasm surrounding its use as a predictor of job performance.

What degree of utility do these global Big Five measures offer for predicting job performance? Overall, it appears that global measures of Conscientiousness can be expected to consistently add a small portion of explained variance in job performance across jobs and across criterion dimensions. In addition, for certain jobs and for certain criterion dimensions, certain other Big Five dimensions will likely add a very small but consistent degree of explained variance. If the global Big Five measure is uncorrelated with the other predictors that are currently used for a job (e.g., personality tends to be uncorrelated with cognitive ability; Day & Silverman, 1989; Rosse, Miller, & Barnes, 1991), then even this small incremental explained variance can, under certain circumstances, make a practically significant contribution to predictive efficiency for a job and perhaps contribute to a reduction in adverse impact (Hattrup, Rock, & Scalia, 1997; Murphy & Shiarella, 1997; Schmitt, Rogers, Chan, Sheppard, & Jennings, 1997; but see Ryan, Ployhart, & Friedel, 1998).

However, if the relevant aspects of the applicant's personality are already partially captured through other selection techniques such as reference checks and interviews, the small potential contribution of a global Big Five measure will likely diminish. In addition, the potentially negative impact of faking on the utility of personality measures (Christiansen, Goffin, Johnston, & Rothstein, 1994; Ellingson, Sackett, & Hough, 1999; Rosse, Stecher, Miller, & Levin, 1998) and the potentially negative applicant reactions (Rosse, Miller, & Stecher, 1994; Steiner & Gilliland, 1996) also raise the question of whether this small addition of explained variance should be interpreted as enthusiastically as it recently has been. Thus, although the low to moderate validity coefficient for Conscientiousness tends to generalize quite well across occupations and job performance criteria and although other Big Five dimensions appear to have meaningful relations with certain criteria or for certain jobs, we do not see evidence that Conscientiousness or any of the other Big Five dimensions should be granted a status similar to that of general cognitive ability for personnel selection purposes (cf. Behling, 1998).

In terms of theory rather than practice, however, we do interpret our findings as indicating a pattern of theoretically meaningful relations between the broad personality dimensions and job performance that should be explored in future research, perhaps using facet scales of the Big Five dimensions. Although the strength of the relations are low to moderate, different personality dimensions appear to affect performance in different types of jobs or along different dimensions. In a relative sense, the Conscientiousness dimension does appear to have the strongest relation to overall job performance. People who describe themselves as hard-working, reliable, organized, and so on do appear to perform somewhat better than do those who believe they are less strong in these characteristics. It is also interesting that Emotional Stability showed a rather stable influence on performance throughout nearly all of our analyses. It appears that being calm, secure, well-adjusted, and low in anxiety has a small but consistent impact on job performance. Agreeableness also gains importance for those jobs that require interpersonal interactions, so that being likeable, cooperative, and good-natured has a small but consistent impact on performance. Finally, being Extraverted appears to influence sales and perhaps managerial jobs, and Openness to Experience appears to affect performance in customer service jobs. Although these theoretically meaningful relations are rather low in magnitude at

the broad dimension level of the Big Five, the magnitude of these correlations might be enhanced if the most relevant specific facets of these broad dimensions could be specified.

We suggest, then, that the Big Five framework and the patterns of small to moderate validities for these broad dimensions that have begun to emerge should be used in future research to help guide the selection back "downward" toward somewhat narrower personality facets with theoretical links to the performance dimensions under investigation. If a broad, global performance criterion is of interest, perhaps a global Conscientiousness scale will suffice with a moderate level of validity. However, if multiple performance dimensions such as those distinguishing task performance from contextual performance, or perhaps those consistent with other typologies such as that presented by Campbell (1990), will be delineated, then perhaps narrower facets of performance with strong theoretical links to those criteria can be identified and used individually or in combination to enhance their criterion-related validity.

We also note that the formation of optimal composites may involve grouping facets from across the five broad dimensions. For example, combining selected facets of Conscientiousness, Emotional Stability, and Agreeableness may optimize the prediction of an interpersonal facilitation criterion. The circumplex models of the Big Five presented by Hofstee, de Raad, and Goldberg (1992) and Johnson and Ostendorf (1993) could also prove useful in this regard by guiding the formation of predictor scales that simultaneously represent aspects of two dimensions, in a sense falling between two of the broader dimensions. This is a hypothesis that deserves consideration in future research.

In addition to further exploration of more specific links between dimensions of personality and job performance, research that more completely delineates the nomological network connecting personality to job performance is needed. Much of the research to date has taken a very practical perspective, focusing on the bivariate correlation between personality and performance. However, if we are to truly understand the relationship between personality and job performance, we must move beyond this bivariate relationship and toward specifying the intervening variables that link these domains. The Conscientiousness trait, for example, is often discussed in a manner that assumes it has motivational implications. Motivational variables, then, should be examined more extensively as intervening variables in a multivariate model. Some research has, in fact, found Conscientiousness to influence performance through its effects on such motivational variables as performance expectancies, self-efficacy, and goal setting (Barrick, Mount, & Strauss, 1993; Gellatly, 1996; Martocchio & Judge, 1997). Better explication of this nomological network for different conceptualizations of the personality domain (e.g., narrower facets or Big Five dimensions other than Conscientiousness) and for different dimensions of job performance may aid in a better understanding of how personality affects job performance. With better identification of intervening variables, the total effects of personality on job performance may emerge more strongly than the simple bivariate correlation coefficient has demonstrated.

Finally, for the sake of understanding the impact of personality on job performance, it would also be interesting to explore these relations using alternative measurement methods. Mount, Barrick, and Strauss (1994), for example, presented some evidence that supervisor, coworker, and customer ratings of personality consis-

tently had equivalent or higher levels of criterion-related validity in comparison with employees' self-reports. Although the practice of using rating sources other than oneself is not likely to be adopted in personnel selection practice, such alternative measurement methods could help gain a better understanding of the aspects of personality that affect performance.

Limitations

At least two limitations of the current meta-analysis should be pointed out. First, several of our moderator analyses were based on a relatively small number of correlations, especially for the managerial occupation category and for training proficiency. This renders any conclusions based on these moderator categories tenuous. It is especially unfortunate that so little research has been published using managers, because clarification on the impact of the Big Five for this occupation would be beneficial for selection research and practice.

Second, our categorization of job performance dimensions into task performance, job dedication, and interpersonal facilitation could be criticized on the same grounds that we used to criticize the earlier meta-analyses for their categorization of personality measures into the Big Five. Perhaps as the body of research on dimensions of job performance develops, an assessment of the predictability of these or other performance dimensions using a priori measures of the relevant dimensions can be undertaken so as to avoid any problems with this type of classification procedure. Given that we found few effects of such potential for classification errors on the resulting validities for the Big Five, we may be able to predict that such an analysis would yield results that differ little from those we obtained in this study. However, this is an empirical question that would need to be addressed directly in future research.

Conclusions

In summary, the present meta-analysis provides a review of the criterion-related validities of the Big Five personality dimensions, as measured by scales that were developed explicitly according to the five-factor model. Although we have interpreted this evidence of the criterion-related validity of Conscientiousness somewhat less optimistically than many researchers have tended to do in recent years, we nevertheless suggest that the potential exists for improving the validity of personality predictors. We encourage future research aimed at theory-based matching of personality constructs and dimensions of job performance, perhaps using composites of narrower Big Five facets. We also encourage research aimed at building a more extensive multivariate model of the personality–job performance relation.

We conclude that global measures of the Conscientiousness dimension have a rather moderate impact on performance, although this validity does appear rather stable and generalizable across occupations and criteria. Although they are less generalizable, we also conclude that personality traits other than Conscientiousness are nearly equally important for certain occupations and criteria. Our hope is that the results of this review encourage realistic expectations about the potential contribution of Conscientiousness measures to selection utility and encourage further

exploration of the impact of personality variables on job performance beyond the global Conscientiousness dimension.

References

References marked with an asterisk indicate studies that contributed a validity coefficient to the meta-analysis.

Barrick, M. R., & Moutti, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1–26.

*Barrick, M. R., & Mount, M. K. (1993). Autonomy as a moderator of the relationships between the Big Five personality dimensions and job performance. *Journal of Applied Psychology, 78*, 111–118.

*Barrick, M. R., & Monnt, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology, 81*, 261–272.

*Barrick, M. R., Mount, M. K., & Strauss, J. P. (1993). Conscientiousness and performance of sales representatives: Test of the mediating effects of goal setting. *Journal of Applied Psychology, 78*, 715–722.

*Barrick, M. R., Mount, M. K., & Strauss, J. P. (1994). Antecedents of involuntary turnover due to a reduction in force. *Personnel Psychology, 47*, 515–535.

Behling, O. (1998). Employee selection: Will intelligence and conscientiousness do the job? *Academy of Management Executive, 12*, 77–86.

Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 71–98). San Francisco: Jossey-Bass.

Borman, W. C., & Motowidlo, S. J. (1997). Task performance and contextual performance: The meaning for personnel selection research. *Human Performance, 10*, 99–109.

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 687–732). Palo Alto, CA: Consnlting Psychologists Press.

*Cellar, D. F., DeGrendel, D. J. D., Klawsky, J. D., & Miller, M. L. (1996). The validity of personality, service orientation, and reading comprehension measures as predictors of flight attendant training performance. *Journal of Business and Psychology, 11*, 43–55.

*Cellar, D. F., Miller, M. L., Doverspike, D. D., & Klawsky, J. D. (1996). Comparison of factor structures and criterion related validity coefficients for two measures of personality based on the five factor model. *Journal of Applied Psychology, 81*, 694–704.

Christiansen, N. D., Goffin, R. D., Johnston, N. G., & Rothstein, M. G. (1994). Correcting the 16PF for faking: Effects on criterion-related validity and individual hiring decisions. *Personnel Psychology, 47*, 847–860.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual.* Odessa, FL: Psychological Assessment Resonrces.

*Crant, J. M. (1995). The Proactive Personality Scale and objective job performance among real estate agents. *Journal of Applied Psychology, 80*, 532–537.

Day, D. V., & Silverman, S. B. (1989). Personality and job performance: Evidence of incrementental validity. *Personnel Psychology, 42*, 25–36.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417–440.

Ellingson, J. E., Sacken, P. R., & Hough, L. M. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal of Applied Psychology, 84*, 155–166.

Gellatly, I. R. (1996). Conscientiousness and task performance: Test of a cognitive process model. *Journal of Applied Psychology, 81*, 474–482.

Goldberg, L. R. (1992). The development of markers of the big-five factor structure. *Psychological Assessment, 4*, 26–42.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48*, 26–34.

Guion, R. M., & Gottier, R. F. (1965). Validity of personality measures in personnel selection. *Personnel Psychology, 18*, 135–164.

Hattrup, K., Rock, J., & Scalia, C. (1997). The effects of varying conceptualizations of job performance on adverse impact, minority hiring, and predicted performance. *Journal of Applied Psychology, 82*, 656–664.

*Hayes, T. L., Roehm, H. A., & Castellano, J. P. (1994). Personality correlates of success in total quality manufacturing. *Journal of Business and Psychology, 8*, 397–411.

Hofstee, W. K. B., de Raad, B., & Goldberg, L. R. (1992). Integration of the Big Five and circumplex approaches to trait strncture. *Journal of Personality and Social Psychology, 63*, 146–163.

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

Hogan, R., Hogan, J., & Roberts, B. W. (1996). Personality measurement and employment decisions: Questions and answers. *American Psychologist, 51*, 469–477.

Hogan, J., & Ones, D. S. (1997). Conscientiousness and integrity at work. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 849–870). San Diego, CA: Academic Press.

*Hogan, J., Rybicki, S. L., Motowidlo, S. J., & Borman, W. C. (1998). Relations between contextual performance, personality, and occupational advancement. *Human Performance, 11*, 189–207.

Hough, L. M. (1992). The "Big Five" personality variables—construct confusion: Description versus prediction. *Human Performance, 5*, 139–155.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.

Hunter, J. E., Schmidt, F. L., & Judiesch, M. K. (1990). Individual differences in output variability as a function of job complexity. *Journal of Applied Psychology, 75*, 28–43.

Iaffaldano, M. T., & Muchinsky, P. M. (1985). Job satisfaction and job performance: A meta-analysis. *Psychological Bulletin, 97*, 251–273.

*Jacobs, R. R., Conte, J. M., Day, D. V., Silva, J. M., & Harris, R. (1996). Selecting bus drivers: Multiple predictors, multiple perspectives on validity, and multiple estimates of utility. *Human Performance, 9*, 199–217.

John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66–100). New York: Guilford.

Johnson, J. A., & Ostendorf, F. (1993). Clarification of the five-factor model with the abridged Big Five dimensional circumplex. *Journal of Personality and Social Psychology, 65*, 563–576.

*Kolz, A. R., Cardillo, E. P., & Pena. S. A. (1998, April). *Cognitive ability and personality as predictors of performance and counterproductive behavior.* Paper presented at the annual conference of the Society for Industrial and Organizational Psychology, Dallas, TX.

Law, K. S., Schmidt, F. L., & Hunter, J. E. (1994). A test of two refinements in procedures for meta-analysis. *Journal of Applied Psychology, 79*, 978–986.

*Lyne, R., Sinclair, R. R., & Gerhold, C. (1997, April). *Personality and job performance: Matching predictor and criterion domains.* Paper presented at the annual conference of the Society for Industrial and Organizational Psychology, St. Louis, MO.

*Mabon, H. (1998). Utility aspects of personality and performance. *Human Performance, 11*, 289–304.

*Martocchio, J. J., & Judge, T. A. (1997). Relationship between conscientiousness and learning in employee training: Mediating inlluences of

self-deception and self-efficacy. *Journal of Applied Psychology, 82,* 764–773.

Motowidlo, S. J., & Van Scotter, J. R. (1994). Evidence that task performance should be distinguished from contextual performance. *Journal of Applied Psychology, 79,* 475–480.

Mount, M. K., & Barrick, M. R. (1995). The Big Five personality dimensions: Implications for research and practice in human resources management. In K. M. Rowland & G. Ferris (Eds.), *Research in personnel and human resources management* (Vol. 13, pp. 153–200). Greenwich, CT: JAI Press.

*Mount, M. K., Barrick, M. R., & Stewart, G. L. (1998). Five-factor model of personality and performance in jobs involving interpersonal interactions. *Human Performance, 11,* 145–165.

*Mount, M. K., Barrick, M. R., & Strauss, J. P. (1994). Validity of observer ratings of the Big Five personality factors. *Journal of Applied Psychology, 79,* 272–280.

Murphy, K. R., & Shiarella, H. (1997). Implications of the multidimensional nature of job performance for the validity of selection tests: Multivariate frameworks for studying test validity. *Personnel Psychology, 50,* 823–854.

Ones, D. S., Mount, M. K., Barrick, M. R., & Hunter, J. E. (1994). Personality and job performance: A critique of the Tett, Jackson, and Rothstein (1991) meta-analysis. *Personnel Psychology, 47,* 147–156.

*Piedmont, R. L., & Weinstein, H. P. (1994). Predicting supervisor ratings of job performance using the NEO Personality Inventory. *Journal of Psychology, 128,* 255–265.

*Robie, C., & Ryan, A. M. (1998, April). *Effects of non-linearity and heteroscedasticity on the validity of conscientiousness in predicting overall job performance.* Paper presented at the annual conference of the Society for Industrial and Organizational Psychology, Dallas, TX.

*Rose, R. M., Fogg, L. F., Helmreich, R. L., & McFadden, T. J. (1994). Psychological predictors of astronaut effectiveness. *Aviation, Space, and Environmental Medicine, 65,* 910–915.

Rosse, J. G., Miller, H. E., & Barnes, L. K. (1991). Combining personality and cognitive ability predictors for hiring service-oriented employees. *Journal of Business and Psychology, 5,* 431–445.

Rosse, J. G., Miller, J. L., & Stecher, M. D. (1994). A field study of job applicants' reactions to personality and cognitive ability testing. *Journal of Applied Psychology, 79,* 987–992.

Rosse, J. G., Stecher, M. D., Miller, J. L., & Levin, R. A. (1998). The impact of response distortion on preemployment personality testing and hiring decisions. *Journal of Applied Psychology, 83,* 634–644.

Rothstein, H. R. (1990). Interrater reliability of job performance ratings: Growth to asymptote level with increasing opportunity to observe. *Journal of Applied Psychology, 75,* 322–327.

Ryan, A. M., Ployhart, R. E., & Friedel, L. A. (1998). Using personality testing to reduce adverse impact: A cautionary note. *Journal of Applied Psychology, 83,* 298–307.

*Rybicki, S. L., & Klippel, D. C. (1997, April). *Exploring the impact of personality syndromes on job performance.* Poster session presented at the 12th annual convention for the Society of Industrial and Organizational Psychology, St. Louis, MO.

Salgado, J. F. (1997). The five factor model of personality and job performance in the European community. *Journal of Applied Psychology, 82,* 30–43.

*Salgado, J. F., & Rumbo, A. (1997). Personality and job performance in financial services managers. *International Journal of Selection and Assessment, 5,* 91–99.

*Schmit, M. J., Motowidlo, S. J., Degroot, T. G., Cross, T. C., & Kiker, D. S. (1996, April). Explaining the relationship between personality and job performance. In J. M. Collins (Chair), *Personality predictors of job performance: Controversial issues.* Symposium conducted at the annual convention for the Society of Industrial and Organizational Psychology, San Diego, CA.

Schmitt, N. W., Gooding, R. Z., Noe, R. A., & Kirsch, M. (1984). Meta-analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology, 37,* 407–422.

Schmitt, N., Rogers, W., Chan, D., Sheppard, L., & Jennings, D. (1997). Adverse impact and predictive efficiency of various predictor combinations. *Journal of Applied Psychology, 82,* 719–730.

Steiner, D. D., & Gilliland, S. W. (1996). Fairness reactions to personnel selection techniques in France and the United States. *Journal of Applied Psychology, 81,* 134–141.

*Stewart, G. L. (1996). Reward structure as a moderator of the relationship between extraversion and sales performance. *Journal of Applied Psychology, 81,* 619–627.

*Stewart, G. L. (1997, August). Applicants versus incumbents: Assessing the impact of validation design on personality research. In N. Schmitt (Chair), *Relating personality to job attitudes and job performance.* Symposium conducted at the 1997 Annual Meeting of the Academy of Management, Boston.

*Stewart, G. L., & Carson, K. P. (1995). Personality dimensions and domains of service performance: A field investigation. *Journal of Business and Psychology, 9,* 365–378.

*Stewart, G. L., Carson, K. P., & Cardy, R. L. (1996). The joint effects of conscientiousness and self-leadership training on employee self-directed behavior in a service setting. *Personnel Psychology, 49,* 143–164.

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology, 44,* 703–742.

Tett, R. P., Jackson, D. N., Rothstein, M., & Reddon, J. R. (1994). Meta-analysis of personality–job performance relations: A reply to Ones, Mount, Barrick, and Hunter (1994). *Personnel Psychology, 47,* 157–172.

*Van Scotter, J. R., & Motowidlo, S. J. (1996). Interpersonal facilitation and job dedication as separate facets of contextual performance. *Journal of Applied Psychology, 81,* 525–531.

Whitener, E. M. (1990). Confusion of confidence intervals and credibility intervals in meta-analysis. *Journal of Applied Psychology, 75,* 315–321.