PERSONNEL PSYCHOLOGY
2008, 61, 827–868

INCREASED ACCURACY FOR RANGE RESTRICTION CORRECTIONS: IMPLICATIONS FOR THE ROLE OF PERSONALITY AND GENERAL MENTAL ABILITY IN JOB AND TRAINING PERFORMANCE

FRANK L. SCHMIDT, JONATHAN A. SHAFFER, AND IN-SUE OH
Department of Management and Organizations, Tippie College of Business
University of Iowa

A recently developed procedure produces substantial improvements in the accuracy of corrections for range restriction and reveals that predictive validities of employment selection methods are higher than previously thought. This study applied this procedure to meta-analytic validities of 2 personality measures (Conscientiousness and Emotional Stability) and general cognitive ability measures. Results show that the increases in validity estimates are greater for cognitive ability than for personality, reducing the relative validity of personality. In addition, the incremental validity of these personality measures over that provided by cognitive ability is smaller than previously estimated. Path analyses based on the more accurate data reveal relatively smaller path coefficients from these personality traits to job and training performance. These findings have implications for theories of job performance and for the practical value in selection of personality measures relative to cognitive ability measures.

A major goal in any research area is the development of theory, or explanations, for the phenomena studied. Theories are of necessity built on a foundation of facts that usually consist of relationships that are considered to be well established and that any proposed theory must incorporate and explain. A central construct of interest in industrial-organizational psychology is that of job performance, and a major theoretical task is the development of theories of job performance (Campbell, 1990; Campbell et al., 1990; Campbell, McCloy, Oppler, & Sager, 1993; Hunter & Schmidt, 1996; Ree, Earles, & Teachout, 1994; Viswesvaran, Schmidt, & Ones, 2005). Constructs that predict job performance may be incorporated into such theories as hypothesized causes of job performance. In recent years, personality trait constructs as assessed in the Big Five model of personality (Digman, 1990; Goldberg, 1990; McCrae & Costa, 1987, 1989) and the construct of general mental ability (GMA; Brown, Le, & Schmidt, 2006;

Correspondence and requests for reprints should be addressed to Jonathan A. Shaffer, Department of Management and Organizations, Tippie College of Business, University of Iowa, W252 John Pappajohn Business Building, Iowa City, IA 52242; jonathan-shaffer@uiowa.edu.

Jensen, 1998; Ree & Earles, 1991, 1992; Ree et al., 1994; Schmidt, 2002; Schmidt & Hunter, 1992, 1998, 2004) have received considerable emphasis in theory development (e.g., Barrick & Ryan, 2003; Judge, Higgins, Thoresen, & Barrick, 1999). At the same time, other measured variables, such as biographical data measures, employment interview scores, and assessment center scores, have received less emphasis in theory construction because it is less clear what constructs are assessed by these measures (Schmidt & Hunter, 1998).

The results of meta-analyses suggest that two of the Big Five personality traits—Conscientiousness and Emotional Stability—may be valid for predicting performance on most or all jobs (Barrick & Mount, 1991; Hurtz & Donovan, 2000; Mount & Barrick, 1995; Salgado, 1997, 2003). In this study, we present findings for all of the Big Five personality traits, but our primary presentation and interpretation focuses on Conscientiousness (Con) and Emotional Stability (ES). Meta-analytic findings also indicate that GMA is valid for predicting performance on essentially all jobs (Bertua, Anderson, & Salgado, 2005; Hunter, 1980, 1983, 1985, 1986; Hunter & Hunter, 1984; Salgado, Anderson, Moscoso, Bertua, & de Fruyt, 2003; Salgado, Anderson, Moscoso, Bertua, de Fruyt, & Rolland, 2003; Schmidt & Hunter, 1998). For theory development purposes, an accurate calibration of the absolute and relative sizes of these relationships is important, especially in light of the fact that these relationships may be hypothesized as causal in nature within theories of job performance. Accurate calibrations of the relationships of these variables with job performance constitute the essential facts that are the "building blocks" of theory. In addition, there are also important practical implications in terms of the utility (practical value) of employment selection systems that incorporate personality and mental ability measures (cf. Schmidt, Hunter, McKenzie, & Muldrow, 1979).

General Context and Purpose of the Study

As explained in more detail below, Hunter, Schmidt, and Le (2006) and Schmidt, Oh, and Le (2006) have recently presented a new procedure for correcting for the downward bias in predictor-job performance relationships caused by range restriction, and this procedure has been shown analytically and via computer simulation to be more accurate than the range restriction correction procedure typically used to date. Schmidt et al. (2006) showed that this more accurate procedure gives estimates of validity for mental ability measures that are substantially higher (30% higher for job performance and 42% higher for training performance) than those presented in past meta-analyses in the literature. As of yet, this new procedure has not been applied to validities for personality measures. The purpose of this study is to do this and to thereby more accurately calibrate the validities of measures of the personality traits of Con and ES, both absolutely and relative to the validity of GMA measures. The databases we use in this connection are those presented by meta-analyses in the literature for the two personality traits and GMA. In each case we correct the mean observed validities in these meta-analyses first for direct range restriction (that is, using the older method of correction) and then for indirect range restriction using the new procedure described above and discussed in detail below. By holding all other variables (e.g., validity data, artifact data, etc.) constant, we can observe the effect on validity estimates of moving from the older to the newer range correction procedure (as was done in Schmidt et al., 2006).

Background on Range Restriction Correction Methods

In many research situations, such as educational and employment selection, researchers have data only from a restricted population and yet must attempt to estimate parameters of the unrestricted population (cf. Sackett & Yang, 2000). For example, the validity of the medical college aptitude test (MCAT) for predicting performance in medical school can only be estimated using samples of students admitted to medical college (the restricted sample). However, the goal is to estimate the validity of the MCAT when used in the population of *applicants* to medical school. Because of range restriction, the population of admitted students typically has higher mean MCAT scores and a smaller standard deviation (*SD*) of scores. To estimate the validity in the applicant population from the observed validity in the "incumbent" population of admitted students, one must correct for the effects of range restriction on MCAT scores. In such a situation, if applicants have been selected directly on test scores top down, we have what is called *direct* or *explicit* range restriction. On the other hand, if students have been selected on some other variable that is correlated with MCAT scores (such as a composite of undergraduate grade point average and letters of recommendation), then the range restriction is said to be indirect.

Based on the earlier work of Pearson (1903), Thorndike (1949) presented corrections for direct and indirect univariate range restriction; that is, equations for correcting for range restriction when restriction has occurred on only one variable. His case II correction for direct range restriction is widely used, as described later. His case III correction for indirect range restriction produced by direct restriction on a third, known variable can rarely be used because its use requires considerable information on the third variable and in most research this information is unknown, as explained below.

As noted nearly 6 decades ago by Thorndike (1949, p. 175), most range restriction in personnel selection and other areas is indirect. It is very rare for individuals to be selected and rejected top down on any single measure in employment, educational, or other selection, as is required in the definition of direct range restriction. For example, in the general aptitude test battery (GATB) database of the U.S. Department of Labor (Hunter & Hunter, 1984), none of the 515 studies were characterized by direct range restriction; in all of them range restriction was indirect (Hunter et al., 2006). All were concurrent validation studies and in none of the studies had GATB scores been used earlier to select the incumbents. Typically, a number of variables are used in selection, and selection decisions are based on a (weighted) sum or composite of these variables. If selection is top down, it produces direct range restriction on the composite. However, typically at least some of the variables in the composite are subjective judgments (e.g., unquantified interview impressions, impressions based on letters of recommendation, etc.) and are never coded explicitly as scores. As a result, the degree of direct range restriction on the composite (i.e., the ratio of the restricted to the unrestricted *SD*) cannot be assessed (because the *SD*s of the composites cannot be computed).

In theory, a method for correcting for indirect range restriction has long been available, namely Thorndike's (1949) case III correction equation for indirect range restriction. This equation can be written as follows (Thorndike, 1949, p. 174):

$$R_{xy} = \frac{r_{xy} + \left(\frac{1}{u_z^2} - 1\right) r_{xz} r_{yx}}{\sqrt{\left[\left(\frac{1}{u_z^2} - 1\right) r_{xz}^2 + 1\right]\left[\left(\frac{1}{u_z^2} - 1\right) r_{yz}^2 + 1\right]}}, \qquad (1)$$

where R_{xy} = the validity corrected for range restriction; r_{xy} = the restricted group observed validity; r_{xz} = the restricted group observed correlation between the predictor x and the third variable z; r_{yz} = the restricted group correlation between the criterion y and the third variable z; and $u_z = s_z/S_z$, where s_z and S_z are the restricted and unrestricted *SD*s of z, respectively.

In practice, this equation can rarely be used. This equation assumes that there is indirect restriction on the measure of interest (x) as a result of direct restriction on a third variable (z). For example, people have been selected top down on the single test z (e.g., a measure of GMA), which is correlated with the predictor of interest x (e.g., a work sample test), producing indirect restriction on x. Use of the case III correction requires that (a) direct top down selection has been based solely on z and has not been affected by any other information (unlikely to be true); and (b) the following data are known: r_{xz}, r_{yz}, r_{xy}, s_z and S_z. These conditions are almost never met. Even if the first condition was met, the information required in the second condition would rarely be available. Hence, the case III correction is almost never used (Hunter & Schmidt, 2004, ch. 5; Hunter et al., 2006; Linn, Harnisch, & Dunbar, 1981). Instead, Thorndike's familiar case II equation, which assumes direct selection on x, has been almost universally used, both in individual studies (Cohen & Cohen, 1983, p. 70) and in meta-analysis methods (Callender & Osburn, 1980; Hunter & Schmidt, 1990; Raju & Burke, 1983). The case II equation can be written (Thorndike, 1949, p. 173):

$$R_{xy} = \frac{\left(\frac{1}{u_x}\right) r_{xy}}{\sqrt{\left[\left(\frac{1}{u_x^2}\right) - 1\right] r_{xy}^2 + 1}}, \tag{2}$$

where R_{xy} = the validity corrected for range restriction, r_{xy} = the observed validity in the restricted (incumbent) group, and $u_x = s_x/S_x$, where s_x and S_x are the restricted and unrestricted *SD*s of x, respectively.

It has been shown that use of the case II correction when range restriction is indirect leads to substantial undercorrection for the effects of range restriction and thus underestimates the true (or operational) validity (Hunter & Schmidt, 2004, ch. 5; Hunter et al., 2006; Linn et al., 1981; Schmidt, Hunter, Pearlman, & Hirsh, 1985, p. 751). Hunter et al. (2006) and Hunter and Schmidt (2004) presented a new procedure, called "Case IV," for correcting for indirect range restriction that does not depend on conditions a and b required for use of the case III correction. This procedure can be used in the common case in which direct selection has taken place on some unmeasured (and typically unknown) composite of variables (subjective or objective). The procedure assumes that range restriction on this unmeasured composite creates indirect restriction only on x_t (the predictor construct) and has no effect on y_t (the criterion construct) except through x. If this assumption is not violated, then case IV is mathematically identical to Thorndike's case III and is very accurate (Le, 2003; Le & Schmidt, 2006). Extensive computer simulation has shown that even if this assumption is violated, the new procedure is still quite accurate (although less so) and is considerably more accurate than the typically used case II correction equation. When this assumption is violated, the case IV procedure generally undercorrects for the effects of range restriction but much less so than the case II equation (Le, 2003; Le & Schmidt, 2006).

So under all conditions of indirect range restriction the new procedure is more accurate than the case II equation.

A key difference between direct and indirect range restriction is that in the direct case, range restriction occurs on observed scores (because selection is based on observed scores), although in the indirect case, observed scores are not used in selection. As a result, in the indirect case range restriction occurs on true scores rather than on observed scores (Hunter & Schmidt, 2004, ch. 5; Hunter et al., 2006). In the new correction procedure for indirect range restriction, one must first estimate the degree of range restriction on the true scores t underlying the observed score x from the degree of range restriction on the observed scores x. As in case II, the latter is defined as $u_x = s_x/S_x$; that is, the ratio of the restricted to the unrestricted observed *SD*s of x. By contrast, the quantity $u_t = s_t/S_t$ is the degree of range restriction on the true scores t underlying observed scores x. The equation that gives u_t (Hunter & Schmidt, 2004, Equations 3.16 and 5.31; Hunter et al., 2006, Equation 22) is

$$u_t = \sqrt{\frac{u_x^2 - \left(1 - r_{xx_a}\right)}{r_{xx_a}}}, \tag{3}$$

where u_t and u_x are as defined above, and r_{xx_a} is the reliability of the predictor in the unrestricted group (i.e., applicant group). The equation that corrects for indirect range restriction is then

$$R_{x_t y_t} = \frac{\left(\frac{1}{u_t}\right) r_{x_t y_t}}{\sqrt{\left[\left(\frac{1}{u_t^2}\right) - 1\right] r_{x_t y_t}^2 + 1}}, \tag{4}$$

where $R_{x_t y_t}$ = the correlation between x and y corrected both for range restriction and for measurement error in both x and y, and $r_{x_t y_t}$ is the restricted validity corrected for measurement error in both variables using restricted group (i.e., incumbent group) reliabilities. The estimate of operational validity is then $\sqrt{r_{xx_a}} R_{x_t y_t}$. This final step is necessary because operational validity estimates are not corrected for measurement error in the predictor. It should be noted that this procedure is not accurate unless the observed validity is corrected for measurement error in both variables prior to the correction for range restriction. An explanation for why this is the case, a detailed mathematical derivation of this equation, and a full discussion of the procedure are presented in Hunter and Schmidt (2004; ch. 5) and Hunter et al. (2006).

Method

Because they are more precise and stable than the findings of small sample individual studies, we sought to base our analyses on meta-analytic findings or primary studies based on a large sample size (e.g., Project A; McHenry, Hough, Toquam, Hanson, & Ashworth, 1990). We searched the literature for published meta-analyses or large sample primary studies of the validity of GMA measures, Con measures, and ES measures for predicting overall job performance and performance in job training programs. The initial search process was conducted using several online databases as of December 2006 (PsychInfo, PsycARTICLES, EBSCO, Doctoral Dissertation Database) using relevant keywords. We also searched relevant literature manually using references in the relevant books (i.e., Jensen, 1998) and articles. Our search covered the period from 1977 to 2006 and included the following journals: *Journal of Applied Psychology, Personnel Psychology, Psychological Bulletin, International Journal of Selection and Assessment, Human Performance, Organizational Behavior and Human Decision Processes, Journal of Vocational Behavior, Journal of Organizational Behavior, Journal of Occupational Psychology, Journal of Occupational and Organizational Psychology, Journal of Management,* and *Academy of Management Journal.*

We searched for published meta-analyses and large sample studies that included a variety of jobs (i.e., were not limited to a specific job), included a minimum of at least five studies, were not based on the practically same data sets as other meta-analyses, and reported the psychometric information (i.e., criterion and predictor reliabilities and range restriction ratios) necessary for reanalysis. A major goal of the study was to compare validity estimates resulting from use of the two methods of correcting for range restriction. Hence, it was critical to hold all other factors constant; therefore, we used the same predictor and criterion reliabilities and range restriction ratios in our analysis as had been used in the original studies. In the few cases in which artifact values were not reported, we used empirically based values from other sources in the corrections for both direct and indirect range restriction. This matter is discussed in more detail later in this paper and in Appendix B.

We required that the independent variables be either measures of GMA or measures of Con or ES, as defined in the five-factor model (FFM) discussed earlier, and that criteria be either measures of overall job performance or training performance. The following are examples of decisions made. Meta-analyses based on essentially the same dataset have been reported by (Hunter, 1983, 1985, 1986; and Hunter & Hunter, 1984). We included only Hunter (1986) because it reported the most recent data and included the detailed psychometric information necessary for reanalysis. The Tett, Jackson, and Rothstein (1991) study was excluded because negative signs of validity coefficients were changed to positive prior to the inclusion in their meta-analyses. The meta-analysis by Hogan and Holland (2003) was excluded because the performance criteria they employed were specific dimensions of job performance rather than overall job performance. Appendix A presents more detailed information on the decision rules used and the inclusion/exclusion decisions made.

Using these criteria (described in more detail in Appendix A), we found six meta-analytic estimates for Con using the criterion of overall job performance (Barrick & Mount, 1991; Hough, 1992; Hurtz & Donovan, 2000; Mount & Barrick, 1995; Salgado, 1998; Salgado, 2003) and three using the criterion of performance in training (Barrick & Mount, 1991; Hough, 1992; Mount & Barrick, 1995). For ES, we found five meta-analytic estimates using the criterion of overall job performance (Barrick & Mount, 1991; Hough, 1992; Hurtz & Donovan, 2000; Salgado, 1998; Salgado, 2003) and three using the criterion of training performance (Barrick & Mount, 1991; Hough, 1992; Salgado, 1997). For GMA we found eight meta-analytic estimates of mean validity each for overall job performance and performance in training (Bertua et al., 2005; Hunter, 1986; Salgado, Anderson, Moscoso, Bertua, de Fruyt, & Rolland, 2003a; Salgado, Anderson, Moscoso, Bertua, & de Fruyt, 2003b). All of these meta-analyses (except Hough, 1992) used artifact distribution meta-analysis methods (cf. Hunter & Schmidt, 2004, Ch. 4). These methods have repeatedly been shown via computer simulation methods to produce accurate results (Hunter & Schmidt, 2004; Schmidt & Hunter, 2003), given appropriate estimates of artifact values. (The question of the accuracy of the artifact values used in these studies is addressed later in this paper.) As described in Appendix B, the Hough (1992) study did not correct for any artifacts other than sampling error; in our analysis we corrected these data using the mean artifact values from Barrick and Mount (1991). (See later discussion.)

For each meta-analysis that met our requirements, we recorded the mean observed validity (uncorrected mean correlation) and then corrected this value for criterion unreliability and range restriction, based on the data provided in the meta-analysis. These data included the mean range restriction ratio (u_x) and mean criterion reliability. These corrections are the same as those used in the artifact distribution meta-analysis (Hunter & Schmidt, 2004, Ch. 4). We used the traditional case II procedure described earlier in this paper to apply the correction for direct range restriction (and criterion unreliability). Next, we used the new case IV procedure described earlier to correct for indirect range restriction (and criterion unreliability). No corrections were made for predictor unreliability in these estimates of operational validity.[1] For each meta-analysis, we computed the percentage increase in the mean validity estimate produced by moving from the correction for direct to the correction for indirect range restriction. We also computed the ratio of GMA validity to personality validity for ES and Con under the two range correction procedures. Next, we examined the estimated mean incremental validity for ES and Con, individually and in combination, over that produced by GMA for the two range correction procedures and compared these incremental validities. As stated earlier, our primary focus in the personality domain is on Con and ES, the two consistently valid scales. However, the analyses described above were also conducted for Agreeableness, Extraversion, and Openness to Experience, with these results being presented in Appendix C.

Finally, based on fully corrected (true score) correlations,[2] we applied path analysis in a theoretical causal model and compared the path coefficients in the model across conditions of direct and indirect range restriction corrections to assess changes in absolute and relative causal impact at the construct level as a function of range correction method. The path model was one in which GMA, Con, and ES are all direct causes of performance. This theoretical model of job performance has received considerable attention in the literature (Borman, White, Pulakos, & Oppler, 1991; Campbell, McCloy, Oppler, & Sager, 1993; Schmidt & Hunter, 1992, 1998; cf. also Barrick, Mount, & Judge, 2001 and Mount, Barrick, & Strauss, 1999).

Our original intent was to conduct for contextual performance (CP) the same analyses described here for overall job performance and training performance. However, we found that the validity data in the literature were inadequate to permit this. This was true even when we combined data from studies of CP, organizational citizenship behavior (OCB), and counterproductive work behaviors (CPWB), constructs that are correlated

[1]In examining the meta-analyses in the literature, we found there was sometimes confusion between operational (true) validity and true score correlation. Operational (true) validity is estimated by correcting the observed validity for criterion unreliability and range restriction but not for predictor unreliability. In selection, one predicts actual job performance (not some measure of job performance; hence the correction for criterion unreliability) from observed scores (not from true scores because the employer cannot know true scores; hence the absence of correction for predictor unreliability). The true score correlation differs from the true validity in that it is further corrected for measurement error in the predictor. Hence, it is an overestimate of operational validity in selection. In some cases in the literature the true score correlation is reported as the operational validity. Because of this some of the validity estimates reported in this study appear to be smaller than some in the literature.

[2]Several of the meta-analyses of personality traits reported the average reliability coefficient across the five personality trait measures but not the reliabilities for each trait measure individually. In those cases we contacted the authors and obtained and used the specific reliabilities for the ES and Con scales.

but perhaps not identical (Dalal, 2005). Although we found a reasonable amount of validity data for the two personality traits for CP and OCB (i.e., Organ & Ryan, 1995 for OCB; Salgado, 2002 for CPWB), we found little validity data for GMA. In connection with CPWB, we found only one study that included GMA (Dilchert, Ones, Davis, & Rostow, 2007). In this study, the estimated operational validity for the GMA test for predicting CPWB (measured objectively using personnel records) was −.33, based on correction for direct range restriction. Using the more accurate correction for indirect range restriction, this value is −.57. There are also validity data from Project A for military jobs that are relevant (McHenry et al., 1990). McHenry et al. reported validities for GMA for several criterion measures that are psychologically similar or identical to CP/OCB: (a) effort and leadership ($r = .31$), (b) personal discipline ($r = .16$), and (c) physical fitness and military bearing ($r = .20$). These validity estimates were based on correction for multivariate range restriction (cf. Ree, Carretta, et al., 1994; Sackett & Yang, 2000), a form of correction for indirect range restriction that is nearly identical to the Hunter et al. (2006) correction for indirect range restriction. However, in this study, u_x values were not presented ($u_x = s_x/S_x$) so there is no way to estimate what these validity estimates would be if the correction for direct range restriction were applied. Hence, the McHenry et al. data could not be used in this study.

In connection with the question of CP, it is important to keep in mind that when supervisors' rate the job performance of employees, they weight CP heavily in the ratings (Motowidlo & Van Scotter, 1994; Rotundo & Sackett, 2002). (This is one reason why, even after appropriate corrections for measurement error, supervisor ratings of job performance do not correlate perfectly with job sample measures of performance.) Hence, even when supervisors' ratings are viewed as assessing task performance (i.e., even when the ratings dimensions explicitly refer only to task performance), ratings of overall job performance reflect strong elements of CP (unlike job sample measures of job performance).

McHenry et al. also reported validities for GMA for two measures of general job performance: core technical task proficiency ($r = .63$) and general soldiering proficiency ($r = .65$). As in the case of the other criterion measures, McHenry et al. reported no u_x values. A search of the technical reports related to Project A also failed to uncover these u_x values. Therefore, it was not possible to calculate what the validity estimates would be if they were instead corrected for direct range restriction.[3]

[3]In passing we note a theoretically relevant observation on the McHenry et al. validity estimates for GMA against general job performance measures. These estimates are

Results

Validity Analysis

Tables 1 and 2 present the results for overall job performance and performance in training, respectively. For each meta-analysis, the columns are (1) the number of studies (*k*); (2) the total sample size (*N*); (3) the mean observed (uncorrected) *r*; (4) the mean reliability of the predictor measure in the applicant (unrestricted) sample; (5) the mean reliability of the criterion measure in the incumbent (restricted) group; (6) the mean u_x range restriction ratio ($u_x = s_x/S_x$); (7) the estimated mean operational validity based on correction for direct range restriction (case II); (8) the estimated mean operational validity based on correction for indirect range restriction (case IV); and (9) the percentage increase in the mean validity estimate as one moves from the direct to the indirect range restriction correction.

Because they are estimates of operational validity, the validity estimates in Columns 7 and 8 are not corrected for measurement error in the GMA or personality measures. However, the mean predictor reliabilities presented in Column 4 can be used to correct these estimates, producing the estimated true score correlation (cf. Table 4). True score correlations are useful for theory testing purposes and are used in the theoretical path analysis model presented later (Hunter & Schmidt, 2004; Viswesvaran & Ones, 1995).

We also performed the analysis in Tables 1 and 2 for the remaining FFM personality traits (Agreeableness, Extraversion, and Openness to Experience). These results are shown in Appendix C.

It is clear in Tables 1 and 2 that the increase in the mean validity estimates resulting from use of the case IV range restriction is larger for GMA than for the two personality traits. In both tables, the reason for this differential increase can be seen in Column 6: The mean u_x values are much smaller for the GMA studies (on average, .63 for both job and

---

considerably larger than those reported by others for civilian jobs, and hence, it might have been argued that they were somehow unique to the military and could not be generalized to civilian jobs. For example, Hunter et al. (2006) reported a mean operational validity of .49 (originally from Hunter, 1983) for medium complexity civilian jobs. The McHenry et al. military jobs were also of medium complexity, yet they reported an average validity for GMA of .64. However, as previously mentioned, the McHenry et al. data were corrected for multivariate range restriction (which corrects for indirect range restriction) whereas the most frequently cited civilian studies of the validity of GMA for job performance were corrected for direct range restriction (Hunter, 1983, 1986; Hunter & Hunter, 1984). After reanalyzing the civilian studies using the correction for indirect range restriction, the estimated average validity of GMA for predicting job performance on medium complexity jobs is virtually identical to that of the McHenry data (.66; Hunter et al., 2006). Our study suggests that the McHenry results are not anomalous findings that apply only to military jobs but instead fit nicely into the current mosaic of validity evidence.

TABLE 1
Overall Job Performance Criteria: Operational Validity Estimates for Direct Versus Indirect Range Restriction Corrections

Meta-analysis	(1) k	(2) N	(3) $\bar{r}$	(4) $\bar{r}_{xx_a}$	(5) $\bar{r}_{yy_i}$	(6) $\bar{u}_x$	(7) $\bar{\rho}$ DRR	(8) $\bar{\rho}$ IRR	(9) % Increase
				General mental ability[a]					
Hunter (1986): High complexity jobs	36	2,455	.32	.81	.60	.67	.56[b]	.68	17%
Hunter (1986): Medium complexity jobs	151	12,933	.28	.81	.60	.67	.50[c]	.62	20%
Hunter (1986): Low complexity jobs	201	14,403	.21	.81	.60	.67	.39[c]	.50	23%
Salgado, Anderson, Moscoso, Bertua, de Fruyt, and Rolland (2003): High complexity jobs	14	1,604	.23	.83	.52	.47	.58[d]	.88	34%
Salgado, Anderson, Moscoso, Bertua, de Fruyt, and Rolland (2003): Medium complexity Jobs	43	4,744	.27	.83	.52	.69	.51[d]	.60	16%
Salgado, Anderson, Moscoso, Bertua, de Fruyt, and Rolland (2003): Low complexity jobs	12	864	.25	.83	.52	.67	.48[d]	.59	18%
Salgado, Anderson, Moscoso, Bertua, and de Fruyt (2003)	93	9,554	.29	.83	.52	.62	.58[e]	.71	19%
Bertua et al. (2005)	12	2,469	.22	.85	.52	.60	.47[c]	.61	23%
				Conscientiousness					
Barrick and Mount (1991)	92	12,893	.13	.76	.52	.94	.19[f]	.20	4%
Hough (1992)	169	48,927	.13[g]	.76	.52	.94	.19	.20	4%
Mount and Barrick (1995)	172	31,275	.18	.76	.52	.94	.26[f]	.27	4%
Salgado (1998)	33	4,605	.09	.80	.62	.92	.12[c]	.13	4%
Hurtz and Donovan (2000)	42	7,342	.15	.84	.59	.92	.21[c]	.22	3%
Salgado (2003)	90	19,460	.17	.80	.52	.83	.28	.31	10%

	Emotional Stability								
Barrick and Mount (1991)	87	11,635	.04	.76	.52	.94	.06[f]	.06	4%
Hough (1992)	182	35,148	.09[h]	.76	.52	.94	.13	.14	4%
Salgado (1998)	41	5,203	.08	.80	.62	.92	.11	.12	4%
Hurtz and Donovan (2000)	35	5,027	.09	.86	.59	.92	.13[c]	.13	3%
Salgado (2003)	72	10,786	.09	.80	.52	.81	.15[c]	.18	13%

Note. (1) Number of validity coefficients; (2) total sample size; (3) sample size weighted mean observed validity; (4) mean predictor reliability in the unrestricted population; (5) mean criterion reliability in the restricted population; (6) mean range restriction ratio for observed scores; (7) mean operational (true) validity corrected for direct range restriction; (8) mean operational (true) validity corrected for indirect range restriction; (9) percentage increase in operational validity from applying indirect range restriction correction.

[a]Data from Project A (McHenry et al., 1990) was not included because the artifact values necessary for reanalysis were not reported in the original study.

[b]Mean operational validity reported in the original studies was .58. The value we report here is the mean observed validities corrected for mean criterion reliabilities ($\bar{r}_{yy_i}$) and mean level of range restriction ($\bar{u}_x$). See text for further explanation.

[c]This analysis differs from the originally reported results by .01 due to rounding error.

[d]Mean operational validities reported in the original study were .64, .53, and .51 for high, medium, and low complexity, respectively. The values we report here are mean observed validities corrected for mean criterion reliabilities ($\bar{r}_{yy_i}$) and mean level of range restriction ($\bar{u}_x$). See text for further explanation.

[e]Mean operational validity reported in the original studies was .62. The values we report here are mean observed validities corrected for mean criterion reliabilities ($\bar{r}_{yy_i}$) and mean level of range restriction ($\bar{u}_x$). See text for further explanation.

[f]Barrick and Mount (1991) reported mean true score correlations of .23 for Con and .07 for ES. Mount and Barrick (1995) reported a mean true score correlation of .31 for Con. True score correlations (observed validities corrected for predictor unreliability, criterion unreliability, and range restriction) are of theoretical interest but are not of interest in applied research. Therefore, the values presented here are operational validities (observed validities corrected for criterion unreliability and range restriction but not for predictor unreliability).

[g]The observed validity used in the present analysis is the composite correlation of the observed validities of achievement and dependability presented in Hough (1992; Table 3) computed by using the meta-analytic correlation between the two (.60) from Dudley et al. (2006). See text for further explanation.

[h]The observed validity used in the present analysis was originally reported for adjustment, the label used in that study for ES (Hough, 1992; Table 3).

TABLE 2
Training Performance Criteria: Operational Validity Estimates for Direct Versus Indirect Range Restriction Corrections

Meta-analysis	(1) k	(2) N	(3) $\bar{r}$	(4) $\bar{r}_{xx_a}$	(5) $\bar{r}_{yy_i}$	(6) $\bar{u}_x$	(7) $\bar{\rho}$ DRR	(8) $\bar{\rho}$ IRR	(9) % Increase
	General mental ability								
Hunter (1986): High complexity jobs	24	1,863	.30	.81	.80	.60	.51[a]	.70	27%
Hunter (1986): Medium complexity jobs	54	3,823	.34	.81	.80	.60	.57	.74	24%
Hunter (1986): Low complexity jobs	8	575	.32	.81	.80	.60	.54	.72	25%
Salgado, Anderson, Moscoso, Bertua, de Fruyt, and Rolland (2003): High complexity jobs	13	2,619	.29	.83	.56	.43	.70[b]	.91	24%
Salgado, Anderson, Moscoso, Bertua, de Fruyt, and Rolland (2003): Medium complexity jobs	35	4,304	.29	.83	.56	.67	.53	.64	16%
Salgado, Anderson, Moscoso, Bertua, de Fruyt, and Rolland (2003): Low complexity jobs	21	4,731	.23	.83	.56	.83	.36	.39	8%
Salgado, Anderson, Moscoso, Bertua, and de Fruyt (2003)	97	16,065	.28	.83	.56	.67	.52[b]	.62	17%
Bertua et al. (2005)	53	17,982	.29	.85	.80	.60	.50	.63	22%
	Conscientiousness								
Barrick and Mount (1991)	17	3,585	.13	.76	.52	.94	.19[c]	.20	4%
Hough (1992)	43	5,870	.18[d]	.76	.52	.94	.26	.27	4%
Mount and Barrick (1995)	21	4,106	.17	.76	.52	.94	.25[c]	.26	4%

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Emotional Stability									
Barrick and Mount (1991)	19	3,283	.04	.76	.52	.94	.06[c]	.06	4%
Hough (1992)	69	8,685	.12[e]	.76	.52	.94	.18	.18	4%
Salgado (1997)	6	470	.11	.80	.62	.92	.15[f]	.16	4%

Note. (1) Number of validity coefficients; (2) total sample size; (3) sample size weighted mean observed validity; (4) mean predictor reliability in the unrestricted population; (5) mean criterion reliability in the restricted population; (6) mean range restriction ratio for observed scores; (7) mean operational (true) validity corrected for direct range restriction; (8) mean operational (true) validity corrected for indirect range restriction; (9) percentage increase in operational validity from applying indirect range restriction correction.

[a]This analysis differs from the originally reported results by .01 due to rounding error.

[b]The operational validities reported in the original studies were .74 and .54 for Salgado, Anderson, Moscoso, Bertua, de Fruyt, and Rolland (2003) and Salgado, Anderson, Moscoso, Bertua, and de Fruyt (2003), respectively. The values we report here are mean observed validities corrected for mean level of range restriction ($\bar{u}_x$) and mean criterion reliability ($\bar{r}_{yy_i}$). See text for further explanation.

[c]Barrick and Mount (1991) reported mean true score correlations of .23 for Con and .07 for ES. Mount and Barrick (1995) reported .30. True score correlations (observed validities corrected for predictor unreliability, criterion unreliability, and range restriction) are of theoretical interest but are not of interest in applied research. Therefore, the values presented here are operational validities (observed validities corrected for criterion unreliability and range restriction but not predictor unreliability).

[d]The observed validity used in the present analysis is the composite correlation of the observed validities of achievement and dependability presented in Hough (1992; Table 3) computed by using the meta-analytic correlation between the two (.60) from Dudley et al. (2006). See text for further explanation.

[e]The observed validity used in the present analysis was originally reported for adjustment, the label used in that study for ES (Hough, 1992; Table 3).

[f]The validity reported in the original study (.27) was further corrected for the mean correlation within constructs (see Salgado, 1997, p. 33 for details) and is therefore not appropriate for this analysis. The validity given here is the mean observed validity corrected for criterion unreliability and range restriction.

training performance) than for the personality studies. That is, the level of range restriction on the two personality traits in the incumbent groups is much less than is the case for the GMA measures. In Table 1 all the u_x values for the personality measures are over .90 (They average .91 and .90 for Con and ES, respectively), and in Table 2 all but two are this large (They average .93 for both Con and ES). This indicates that the variability of these scores is nearly as large in the restricted groups (incumbent groups) as it is in applicant pools (suggesting that organizations are selecting and retaining employees based on these traits only to a very limited degree). The result is that the estimates corrected for indirect range restriction are only slightly larger than those corrected for direct range restriction. In general, the more severe the range restriction, the greater is the extent to which the correction for direct range restriction underestimates validity (Hunter et al., 2006). Hence, the increase in validity estimates going from the direct to the indirect range restriction correction is much larger for the GMA measures than for the personality measures.

Table 3 summarizes the findings presented in Tables 1 and 2. The values presented in Table 3 are averages of the values in Tables 1 and 2. Median values (shown in Appendix D) are almost identical to these mean values. The top half of Table 3 presents average values for the criterion of overall job performance. For GMA, moving to the correction for indirect range restriction increases the average validity estimate from .507 to .647, a 27.61% increase. For Con, moving from direct to indirect range restriction corrections increases mean validity from .210 to .220, a 5.71% increase. For ES, the mean validity estimate increases from .117 to .124, a 5.98% increase. Hence, the percentage increase for GMA is 4.8 times larger than that for Con and 4.6 times larger than that for ES.

As a result of these differential percentage increases, the ratios of validities change. When the direct range restriction correction is used (the case in the literature today), the mean validity estimate for GMA (.507) is 2.41 times larger than that for Con (.210). When the more accurate correction of indirect range restriction is used, we see that mean GMA validity (.647) is now 2.91 times as large as that for Con (.222), an increase in this ratio of 20.72%. Hence, both the relative and the absolute importance of GMA have increased substantially. For ES, when the correction for direct range restriction is used (the case in this literature), the mean GMA validity (.507) estimate is 4.43 times as large as that for ES (.117). Applying the more accurate correction for indirect range restriction, we see that the mean GMA validity (.647) estimate is 5.22 times as large as that for ES (.124), an increase of 20.41% in this ratio. Again, the effect of applying the more accurate range correction procedure is to increase the relative importance of GMA in comparison to the personality trait.

Relative sizes of validity estimates can be taken as indices of relative importance (Schmidt et al., 1979). However, we can also assess relative

TABLE 3
Comparison of the Importance of GMA, Con, and ES

	Type of range restriction correction				
	DRR		IRR		
	$\bar{r}_c$	β [f]	$\bar{r}_c$	β [f]	% Increase in $\bar{r}_c$ [a]
Overall job performance criteria					
Average of validities and β weights in Table 1					
GMA	.507	.516	.647	.658	27.61%
Con	.210	.239	.222	.261	5.71%
ES	.117	−.029	.124	−.042	5.98%
Validity and β weights ratios					
GMA : Con	2.41 : 1	2.16 : 1	2.91 : 1	2.52 : 1	20.72%
GMA : ES	4.33 : 1	19.79 : −1	5.22 : 1	15.67 : −1	20.41%
Incremental validity over GMA[b]					
Con	.048[c]		.044[c]		−8.33%
ES	.007[d]		.005[d]		−28.57%
Con & ES	.048[e]		.045[e]		−6.25%
Training performance criterion					
Average of validities and β weights in Table 2					
GMA	.526	.536	.668	.680	27.00%
Con	.235	.265	.245	.286	4.26%
ES	.129	−.030	.135	−.044	4.65%
Validity and β weights ratios					
GMA : Con	2.24 : 1	2.02 : 1	2.73 : 1	2.38 : 1	21.81%
GMA : ES	4.08 : 1	17.87 : −1	4.95 : 1	15.45 : −1	21.35%
Incremental validity over GMA[b]					
Con	.057[c]		.051[c]		−10.53%
ES	.008[d]		.006[d]		−25.00%
Con & ES	.057[e]		.052[e]		−8.77%

[a]Percentage increase in operational validity ($\bar{r}_c$) that results from applying the indirect range restriction correction. Percentage increase does not apply to beta weights.

[b]Operational validities were used in all incremental validity calculations.

[c]The incremental validity of Con over GMA was computed by using the −.03 correlation between GMA and Con reported in Judge, Jackson, Shaw, Scott, and Rich (2007). In order to consistently present calculations based on operational validities, we attenuated the true score correlation between GMA and Con for predictor unreliability in both variables (the true-score correlation reported in the original paper is −.04). The psychometric information needed to compute the operational validity was not reported in the original paper but was provided by Timothy A. Judge.

[d]The incremental validity of ES over GMA was computed by using the .07 correlation between GMA and Con reported in Judge et al. (2007). In order to consistently present calculations based on operational validities, we attenuated the true score correlation between GMA and Con for predictor unreliability in both variables (the true-score correlation reported in the original paper is .09). The psychometric information needed to compute the operational validity was not reported in the original paper but was provided by Timothy A. Judge.

[e]Incremental validity of both Con and ES over GMA is computed using the .46 correlation (the true-score correlation reported in the original paper is .52) between Con and ES reported in Mounte, Barrick, Scullen, and Rounds (2005). Relevant psychometric information was obtained from the method section of the original paper (p. 457).

[f]Our emphasis in this study is on operational validities and incremental validities. However, we have also included standardized regression weights. See discussion in text.

importance by examining the incremental validity produced by the personality traits; that is, the gain in validity from adding the personality trait measure to the GMA measure (e.g., see Schmidt & Hunter, 1998).[4] Estimated average incremental validities for overall job performance are presented in the last three rows in the top half of Table 3, and those for training performance are given in the last three rows of the bottom half of the table. In computing the incremental validity estimates, we used the following meta-analytically based observed correlations from the large study by Judge et al. (2007): $r = -.03$ between GMA and Con and $r = .07$ between GMA and ES. The observed score (i.e., uncorrected) correlation between the two personality measures (.46) was taken from the extensive meta-analysis by Mount, Barrick, Scullen, and Rounds (2005).

Looking at the results for job performance, it can be seen that estimates of incremental validities are somewhat smaller when the more accurate correction for range restriction is used. Incremental validity estimates decline 8.33% for Con, 28.57% for ES, and 6.25% for the combination of both personality traits. Interestingly, the incremental validity from adding ES to the combination of GMA and Con is essentially zero (.045 – .044 = .001).

Turning to the bottom half of Table 3, we see that the overall pattern of results is similar for training performance criteria. For example, the validity estimate for GMA is 2.24 times as large as that for Con under the direct range correction (.526 vs. .235) and becomes 2.73 times as large with the more accurate range correction (.668 vs. .245), an increase in this ratio of 21.81%. For ES, these figures are 4.08 and 4.95, for an increase of 21.35%. Again, the picture is one of decreasing relative importance for the two personality traits as validity estimates become more accurate. As was the case with the criterion of job performance, introducing the more accurate correction for range restriction produces reductions in the estimates of incremental validity for the personality traits. For Con, incremental validity declines from .057 to .051, a 10.53% decrease. For ES, the mean incremental validity estimate declines from .008 to .006.

Path Analysis

Table 4 shows the true score correlations used in the path analyses. The top panel of this table shows the values for the criterion of overall job performance and the bottom panel the values for the criterion of

[4]Our emphasis in this study is on operational validities and incremental validities. However, we have included standardized regression weights in Table 3. See later discussion in text.

TABLE 4
Meta-Analytic True Score Correlations Among Performance, GMA, Con, and ES

	Variable	1	2	3	4
	Job performance				
1	Job performance	1	.712	.251	.139
2	GMA	.558	1	−.040[a]	.090[a]
3	Con	.237	−.040[a]	1	.520[b]
4	ES	.130	.090[a]	.520[b]	1
	Training performance				
1	Training performance	1	.736	.281	.153
2	GMA	.580	1	−.040[a]	.090[a]
3	Con	.270	−.040[a]	1	.520[b]
4	ES	.147	.090[a]	.520[b]	1

Note. All values for job performance and training performance are based on mean meta-analytic correlations corrected for measurement error in the predictor and criterion. Values below (above) the diagonal are further corrected for direct (indirect) range restriction.

[a]This meta-analytic estimate of true-score correlation is reported in Judge et al. (2007).

[b]This (explicitly FFM-based) meta-analytic estimate of true-score correlation is reported in Mount et al. (2005).

training performance. Within each panel, the correlations in the upper part of the matrix are those corrected for indirect range restriction and those in the lower part are those corrected for direct range restriction. The true score correlations between GMA and the two personality traits are from the meta-analysis of Judge et al. (2007). The true score correlation of .52 between Con and ES is taken from the meta-analysis by Mount et al. (2005). The correlations between the three independent variables and the dependent variables (job and training performance) are those from Table 3 after correction for measurement error in the independent variables. The mean validities for GMA in Table 3 are virtually identical to those for medium complexity jobs in Hunter et al. (2006), based on the large U.S. Employment Service database (personality measure validities do not seem to differ by complexity level of job). Hence, the path analysis results presented below can be taken as applying to medium complexity jobs, a category that includes over 60% of all jobs in the U.S. economy (Hunter et al., 2006).

In the first two path analyses, the dependent variable was overall job performance; these path analyses differed only in the type of range restriction applied. The second two path analyses were similar except that the dependent variable was performance in training programs. Figure 1 presents the results for all four path analyses. The path coefficient values above the arrows are those for job performance; the values below the arrows are those for training performance. The values in parentheses are the paths obtained after correction for indirect range restriction; the values

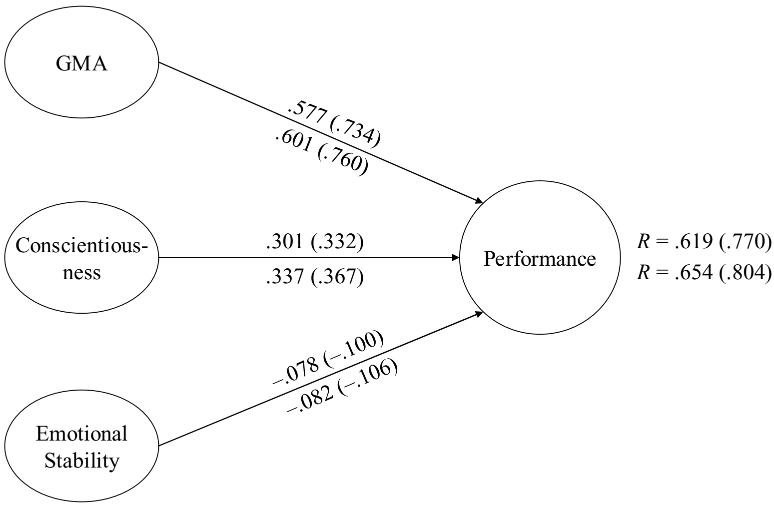

Figure 1: **The Effects of GMA, Con, and ES on Job Performance and Training Performance.**

Note. Upper (lower) values are standardized regression coefficients for job (training) performance. All values are based on mean meta-analytic correlations corrected for measurement error in the predictor and criterion. Values outside (inside) of the parentheses are further corrected for direct (indirect) range restriction.

outside the parentheses are those obtained after correction for direct range restriction.

Considering the first dependent variable, it is clear that estimated absolute and relative causal impact of GMA on job performance increases as we move from the values outside to those inside the parentheses. The path from GMA to job performance increases from .577 to .734, an increase of 27.21%. The increases for Con and ES are much smaller. The path from Con increases from .301 to .332, a 10.30% increase; for ES, the initial path coefficient under the direct range restriction correction is negative, and it becomes more negative when the indirect range restriction correction is used, for a 28.21% decrease. The multiple correlation reflects the extent to which these three traits account for job performance. This value increases substantially when the indirect range correction is introduced, increasing from .619 to .770, a 24.39% increase.[5]

The results for the dependent variable of training performance show a similar pattern. The increase in the path from GMA to training

[5]It is important to remember that these are true score (trait level) multiple correlations and do not reflect the predictive power that would be realized in a selection situation. Multiple correlations based on observed predictor scores would be lower.

performance (.601 to .760) is 26.46%. For Con, this increase 8.90%, and for ES the initially negative path again becomes more negative. The increase in the multiple correlation, from .654 to .804, is 22.94%, again indicating an increase in the explanatory power of the theory when the more accurate correlation estimates are used.

Taken at face value, the negative true score path coefficients seen in Figure 1 for ES suggest that ES is serving as a suppressor variable with respect to one or both of the other two trait measures. Such negative weights for ES can also be seen in Table 3. A suppressor variable is one that when partialled out of another measure increase the other measure's validity or regression weight by "suppressing" or partialling out invalid components of the first measure. Separate analyses (not presented here) showed that the suppressor effect of ES was confined to the Con variable, as would be expected given the relative large true score correlation (.52) between ES and Con and the nearly zero correlations between each personality trait and GMA (See Table 4). As noted earlier, this correlation is from a large sample study (Mount et al., 2005) and is based explicitly on FFM measures. However, it would be desirable in the future to examine other estimates of this relationship. A lower estimate of this correlation might lead to the disappearance of the suppressor effect. But at the present time this is the best available estimate and the only estimate that is based on FFM measures. We would also point out that at least one other study has found suppressor effects in the personality domain. Collins and Schmidt (1997) found larger suppressor effects than the small one found here and showed that the suppressor effects cross-validated in independent samples. This article presents an extended discussion of suppressor variables and their role in measures of personality.

Clearly, the more accurate correlational estimates lead to causal models that suggest a picture of reduced importance for Con and ES relative to GMA in theories of job performance and training performance. The more accurate results also indicate that the ability of the three traits GMA, Con, and ES taken together to account for job and training performance is greater than previously indicated. This increase in the multiple correlations is almost entirely due to the increased importance revealed for GMA as a result of improved accuracy of range restriction corrections.

Discussion and Needed Future Research

In the development of cumulative knowledge and understanding in any scientific area, it is critical to have accurate estimates of the relationships among important theoretical variables because these construct level relationships are the building blocks in the construction of theory and the improvement of applications (Hunter & Schmidt, 2004, pp. 30–31; Viswesvaran & Ones, 1995). Therein lies what is probably the major contribution of this study. Based on data from large meta-analyses in the literature, this study indicates that a method of correcting for the data biases created by range restriction that has been shown to be more accurate than the traditional correction method leads to the conclusion that the two most predictively valid Big Five personality traits, Con and ES, are less important relative to GMA than previously believed. That is, both operational validities and path coefficients in a theoretical model are lower relative to GMA than previously indicated. Incremental validity over GMA is also smaller. The findings are very similar for criteria of overall job performance and performance in job training programs. These conclusions are based on a total of 2,057 studies (1,577 for job performance and 480 for training performance) and 319,288 individuals (241,327 for job performance and 77,961 for training performance) and appear to be robust. However, as noted earlier, we were unable to conduct these analyses for pure measures of CP. Ratings of general job performance do reflect contextual as well as task performance, as stated earlier. However, from the results presented in this study, it is not possible to ascertain the relative or absolute contributions of GMA, Con, and ES to CP per se.

One question that can be raised is how these findings relate to Schneider's (1987) attraction–selection–attrition (ASA) model. The reader should recall that the basic reason why the validity estimates for Con and ES declined relative to those for GMA when the range correction procedure became more accurate is that the initial level of range restriction (u_x value) was much smaller for GMA than for the personality traits. Among incumbents, the *SD* of GMA scores was on average only 63% as large as the *SD* among applicants; for the personality traits this figure was 90%. Schneider (1987) proposed that people who are similar to those already in an organization are more likely to be attracted to that organization (attraction), more likely to be accepted into that organization (selection), and more likely to remain in that organization (low attrition). Schneider (2007) makes clear that the individual difference dimensions the ASA process applies to includes personality *traits*, so one might expect to find as much range restriction on personality traits as on GMA; yet this is apparently not the case. If the job-based data in this study generalize to the organizational level, the suggestion would be that Schneider's ASA process operates more strongly on GMA than on personality. However, this does not imply that the ASA process does not operate on personality traits because it is possible to have substantial mean personality differences between jobs or organizations despite a large u_x value. As discussed in Schmidt, Ones, and Hunter (1992), the relevant principle is a theorem from analysis of variance (ANOVA) that states the total variance is the sum of between-group variance plus average within-group variance. Here between-group variance refers to differences across jobs in mean personality trait levels. It is a fact that is surprising to many that fairly large between-group differences in means produce only small reductions in average within-group variance as compared to total variance. For example, Sackett and Ostgaard (1994), using a large GMA database, found that there were large differences in mean GMA scores of applicants to different jobs (with applicants to higher complexity jobs having higher means and vice versa), yet the average *SD* of GMA scores in applicant pools was only about 8% smaller than in the general workforce norm group. Hence, a mean u_x value of .90 for personality does allow for substantial between-job differences in mean personality trait scores, which if observed would support the ASA model. However, these data do not support the conclusion that the ASA process "produces homogeneity" (Schneider, 2007). The reduction in heterogeneity of about 10% cannot accurately be said to produce homogeneity. Of course, the much smaller u_x values obtained for GMA means that the between-job mean differences in incumbent GMA scores are substantially larger than those for personality scales. However, even with GMA, it would be questionable to argue that the result is homogeneity in GMA within jobs because on average the variability within jobs is still 63% as large as the applicant pool variability.

A model similar in some respects to Schneider's (1987) ASA model is McCormick's gravitational model (McCormick, DeNisi, & Shaw, 1979; McCormick, Jeanneret, & Mecham, 1972). Unlike the ASA model, the gravitational model is focused on jobs. This model postulates that people gravitate toward jobs they are capable of doing well, that survival in a job is an indication of reasonably satisfactory performance, and that poor performance leads to attrition. This model received empirical support in a study by Wilk, Desmarais, and Sackett (1995). This study found that over time people with GMA lower than average for their jobs tended to move to lower complexity jobs and people with GMA higher than average for their jobs tended to move to higher complexity jobs. The gravitation model has been presented and applied in connection with cognitive abilities, either GMA or specific cognitive aptitudes such as spatial, quantitative, verbal, or mechanical, and not in connection with personality traits (McCormick, Jeanneret, & Mecham, 1972). So it does not appear that the present findings on personality contradict the gravitational model. The finding of relatively small u_x values for GMA is consistent with and supportive of the gravitation model.

However, independent of ASA and gravitational models one could raise the question of why personality traits show so little range restriction in comparison to GMA. The first thing to note is that the data on this difference are quite strong: All of the u_x values for personality traits used in the meta-analyses incorporated into this study were empirically determined. (This is also the case in Appendix C for the mean u_x values for Agreeableness, Extraversion, and Openness to Experience.) Often, however, an estimate of the applicant pool *SD* was not available and so a national norm group *SD* was used in its stead. Such norm data typically consist of data pooled across applicants for many different jobs. As noted by Ones and Viswesvaran (2003), this causes the u_x values to be slightly smaller than would be the case using the applicant group *SD* because the norm group *SD* is slightly (about 4% on average) larger than the applicant group *SD*. So if anything, many of the u_x values used in our analyses are too small and therefore overestimate the extent to which there is range restriction on personality. Using data from test manual of the Personal Characteristics Inventory (PCI), a FFM inventory (Mount & Barrick, 2002), we calculated the average u_x across jobs for each of the FFM traits. The norm group *SD* was used as the estimate of the applicant pool *SD*, so these values are slight underestimates. The mean values were as follows: Con (.97), ES (.93), Extraversion (1.12), Agreeableness (.90), and Openness to Experience (.98). These data provide further evidence that, in general, there is little range restriction on personality traits and therefore that organizations do not select or attrite people based on personality to the same extent as they do on GMA. The reader may note that for Extraversion, u_x is greater than 1.00, indicating that on average variation was greater within jobs than in the pooled norm group. The u_x value used in our analysis of each meta-analytic dataset in this study was the average u_x value across the studies included that particular meta-analysis. In connection with sample-based estimates of u_x, we recognize that u_x values can vary from study to study, even after taking into account the fact that much of this variation is sampling error. But it is highly unlikely that *mean* u_x will vary much (for the same trait measure) across different, independent study sets, each made up of a wide variety of jobs. It is important to note that the mean u_x values used in this study are of this nature. This is why the mean u_x values in Tables 1 and 2 do not vary much across meta-analyses of the same trait and for the same criterion (job performance vs. training performance). As shown in Column 6 in Tables 1 and 2, the pattern of findings of lower u_x values for GMA was quite consistent across all data sets. Perhaps the best explanation for this difference is the actual difference in validity between GMA and the personality traits. That is, because the validity of GMA is much larger, organizations are more likely to select, often indirectly, on GMA in hiring, either via standardized GMA measures, interviewing practices that load on GMA, or other means (e.g., educational requirements). Likewise, if retention and termination are heavily conditioned on performance, u_x values for GMA would be further differentially reduced relative to those for personality because the correlation between GMA and performance is larger than the correlation between personality traits and performance. Hence, we believe it should not be unexpected that u_x values among incumbents are larger for personality measures than for GMA measures. The accuracy of u_x values for GMA is discussed in the Limitations section.

Some researchers and I-O psychologists may find it counterintuitive that the most predictively valid personality traits (Con and ES) in the Big Five model could play only a small role in determining job and training performance relative to GMA. We believe it is possible that such intuitions have some merit. It is possible that the results obtained in this study are due in part to the way in which personality has typically been measured in the literature. All of the personality measures used in this study are self-report measures, and this is almost universally the case in the literature. It is a well-known fact that ratings made by multiple raters are more reliable and valid than ratings by a single rater (e.g., see Viswesvaran, Ones, & Schmidt, 1996). Self-report measures have the limitation that it is not possible to have multiple raters (i.e., there is only one self). Mount, Barrick, and Strauss (1994) have reported evidence indicating that when personality is assessed or rated by others acquainted with the target individuals, these measures have higher validity for predicting external criteria. Averaging across multiple raters may increase both reliability and construct validity. The use of this method of measuring personality could lead to research findings indicating that personality—and in particular Con and ES—play a more important role in job performance and training performance than indicated by the findings of this study. This is an area in which more research is needed.

The reliability of self-report personality measures is typically assessed via coefficient alpha and is usually reasonably large (Viswesvaran & Ones, 2000). However, Schmidt, Le, and Ilies (2003) showed that coefficient alpha reliabilities overestimate the reliability of self-report measures of personality (by about .04) because they do not take into account transient measurement error (variations in responses from day to day due to changing moods or other transient psychological states unrelated to the construct being measured). Perhaps more important is the fact that different self-report measures of the same personality traits do not correlate highly. Even when different personality inventories are based on the same Big Five model of personality, corresponding scales typically correlate only about .40 to .60, values much smaller than comparable values for different measures of GMA (Hough, Eaton, Dunnette, Kamp, & McCloy, 1990; Hough, 1992, p. 142, Table 2; Ones et al., 1993, Table 3, p. 129; Schmidt & Hunter, 1999). This suggests that each scale contains large trait-irrelevant specific factors that reduce construct validity (Le, Schmidt, & Putka, in press; Schmidt & Hunter, 1999). It also suggests that from the point of view of construct validity the correlations between corresponding scales across Big Five inventories may better approximate the actual reliabilities of these scales than does coefficient alpha (Le et al., in press; Schmidt & Hunter, 1999). From the point of view of this study, this could mean that the true score correlations for Con and ES used in our path analyses were underestimated because they were undercorrected for measurement error in the personality scales as a result of use of coefficient alpha in the corrections. Use of such alternative reliability estimates could indicate a larger causal role for Con and ES. This effect, although not necessarily large, could be important. For example, if the new estimate of reliability for Con and ES scales were .60, then the estimated true score correlation between each of these personality traits and job and training performance would increase by approximately 15%. If the new reliability estimate were .50, this increase would be approximately 26%.

If correlations between corresponding scales on different Big Five inventories are the best estimates of scale reliability, then reliable measurement of personality traits via self-report might require use of several Big Five inventories, with the corresponding scale scores being averaged across inventories. For example, the measure of Con would be the average across the Con scales from several different inventories. Such measures could be expected to show higher validities than those found in this study for single self-report scales. The estimates of operational validity of self-report scales in Tables 1 through 3 would remain correct but would generalize only to single-scale self-report measures and not to multiscale measures. This is another potentially fruitful topic for future research. Robert Dipboye has suggested other ways in which self-report measures of personality might be made more reliable and valid (see Morgeson, Campion, Dipboye, Hollenbeck, Murphy, & Schmitt, 2007). An ideal approach to measuring personality traits might be to employ both multiple scales for each personality trait *and* multiple external reporters (raters). This approach would address both the validity limitations of single scales and single raters (self-reports).

However, another possibility is that measurement problems in self-report personality measures do not explain all or most of the apparent relative and absolute limited importance of personality traits. A heretofore unexplored theoretical possibility is that it is GMA itself that causes the impact of personality traits on real-life performances to be limited. That is, people may use their GMA to control the expression of their personality tendencies in their behavior. For example, consider an introverted employee in a job that requires meeting and dealing with a wide variety of people, that is, a job in which extroverted behavior is conducive to success. Such a person could use his/her GMA to learn to display such extroverted behaviors, even though such behavior is not his/her natural or spontaneous inclination. This follows from the fact that GMA is an "all purpose tool" that can be used to solve any kind of problem (Gottfredson, 1997; Jensen, 1998), including the problem of how to meet the social requirements of a job. Or consider an executive high in Neuroticism (i.e., low in ES). Given adequate levels of GMA, such a person could learn to control his/her neurotic impulses (e.g., toward hostility or anger) so that there would be little impact on his/her performance or performance ratings. This same process could operate with Con. It would appear that this could be a fertile area for future research exploring the extent to which people do in fact use their GMA to override their personality tendencies in the workplace in the interests of enhanced performance and associated rewards.

One interesting prediction is that the impact of personality on job performance will be found to be smaller for high GMA people but larger for low GMA people. Such a finding would be a major contribution to the understanding of personality, job performance, and the psychology of individual differences. However, it should be noted that this theory hypothesizes an interaction between personality traits and GMA, and studies testing for such interactions have produced mixed results at best. Mount et al. (1999) and Sackett, Gruys, and Ellingson (1998) present reviews of this literature and also report negative findings in their own studies. McClelland and Judd (1993) present a discussion of the methodological reasons why it is difficult to detect interactions. In the absence of an interaction, the implication would be that GMA level does not affect the ability of individuals to control the expression of their personalities.

Limitations

As noted earlier, one limitation of this study is that it was not possible to conduct the analyses with pure measures of CP as the dependent variable. Perhaps future studies can conduct such an analysis when more data become available. It could be argued that a limitation of this paper is that the choices made as to which meta-analyses to include could have affected the results. It is conceivable that other researchers might have made different choices. In response to this possibility, we developed detailed decision rules (described in detail in Appendix A) and implemented them carefully. Further, although the results are not reported here, we looked at how the results changed under different decision rules or inclusion/exclusion of data. We found that the numerical results were similar and that overall conclusions did not change under different decision rules. Hence, these findings appear to be fairly robust.

Another potential limitation stems from the fact that the Hough (1992) meta-analysis of personality traits did not correct for measurement error or range restriction and did not report the information needed to make these corrections. As noted earlier, we used the artifact values from the similar Barrick and Mount (1991) meta-analysis to make these corrections. To assess the impact of the Hough (1992) meta-analysis on our results, we reran our analyses omitting the Hough (1992) data. The numerical results were nearly identical and the conclusions were unchanged. So, this does not appear to be a problem.

In an earlier section, we discussed whether the relatively large u_x values observed for personality measures are accurate or realistic and presented evidence that they are. It is appropriate to similarly examine the u_x values observed and used for GMA and the predictor and criterion reliability values used for both GMA and personality measures in the meta-analyses we included in this study. First, however, we again draw the reader's attention to fact that a major purpose of this study was to compare validity estimates across the two methods of correcting for range restriction. Doing this requires holding all other factors constant, which requires that we use the artifact values from the original meta-analyses in both range restriction corrections (Schmidt et al., 2006). As noted earlier, we did in fact do this and therefore were able to demonstrate the differential impact of the correction for indirect range restriction on GMA and the personality traits of Con and ES.

However, one can also inquire about the absolute accuracy of the validity values produced by the correction for indirect range restriction. For there to be an accuracy problem in this respect, it would have to be the case that the artifact values used in the meta-analyses included in this study were inaccurate in some nontrivial way. A skeptic could raise the question of the accuracy of these artifact values and postulate that unless we can be sure that the artifact values used in these meta-analyses are accurate we cannot be certain that validities corrected for indirect range restriction are indeed more accurate than those corrected for direct range restriction. Appendix B provides considerable information on the artifact values used in the incorporated meta-analyses, including the relevant page numbers in the original articles. These values were almost always empirically derived, whether from the studies in the meta-analysis or from other empirical sources such as test manuals. For example, the reliabilities of tests in the unrestricted group (the applicant pool) were frequently obtained from pooled norm group estimates in test manuals and hence should be quite accurate estimates. The mean values for predictor reliability shown in Tables 1 and 2 are unlikely to be controversial because of their similarity to frequently observed values of the alpha coefficient for GMA and personality scales. Job performance measures are almost invariably ratings of job performance by a single supervisor. The appropriate reliability estimate is therefore interrater reliability (Hunter & Schmidt, 2004; Schmidt, Viswesvaran, & Ones, 2000). Based on independent large data sets, both Rothstein (1990) and Viswesvaran et al. (1996) estimated mean interrater reliability (i.e., the reliability of ratings of overall job performance by one rater) to be .51 to .52. Estimates based on such large data sets are more likely to be accurate than those estimated on the small samples in individual studies. Consequently, these empirically determined values have subsequently been used in a number of meta-analyses, as can be seen in Table 1. The Hunter (1986) meta-analyses were conducted before these estimates were available and, based on a less extensive dataset (see Appendix B), Hunter estimated mean interrater reliability at .60, resulting in a slight undercorrection for criterion reliability. As can be seen in Table 1, meta-analyses using other ratings reliability estimates produced values quite similar to the .52 value discussed above. These facts establish that these earlier published meta-analyses do not suffer from inaccurate estimates of criterion reliability in the job performance domain.

Criterion reliabilities for training performance are usually higher because these measures are usually objective (e.g., objective measures or amount learned in training or objective job sample measures). Hence, as can be seen in Table 2, the mean values for all U.S. meta-analyses of GMA are .80. In Europe, it is more common to assess training performance via ratings, and hence both the Salgado et al. (2003) meta-analyses obtained a lower value (.56). This was also the case in this table for the Salgado (1997) meta-analysis of ES, where the training success measure had a mean reliability of .62. The training criterion measures in Barrick and Mount (1991) and Mount and Barrick (1995) are somewhat unusual for American data in that the authors report that over 90% were based on ratings and not on objective measures of training performance. For this reason, they used the .52 mean value for interrater reliability obtained in the large Viswesvaran et al. (1996) meta-analysis. Hough (1992) also reported heavy use of subjective ratings of training performance, and, as noted earlier, we used the criterion reliability figure from the Barrick–Mount meta-analyses with the Hough (1992) data. Note that if the .52 criterion reliability estimate for the training performance measure in these three meta-analyses is somewhat too low, the result would be a slight overestimate of the validity of Con and ES, both absolutely and relative to the validity of GMA, thus making the results of this study slightly conservative.

The final question is whether the u_x values for GMA in Tables 1 and 2 are accurate. The mean u_x values for the Hunter (1986) meta-analyses are empirically determined averages. In the large U.S. Department of Labor data base used by Hunter, he was able to compute u_x for each study and then take the average across studies, resulting in a mean value of .67 for job performance-based studies ($k = 425$) and .60 for validity studies of training performance ($k = 90$). The figure used as the applicant pool *SD* was the *SD* in the test norm group. Sackett and Ostgaard (1994) showed that this procedure results in little error in u_x estimates (See also Schmidt et al., 1992; and Schmidt, Le, Oh, & Shaffer, 2007). As described in Appendix B, the u_x values used in the other meta-analyses of GMA were computed from the individual studies, obtained from test manuals, or taken from earlier published meta-analyses. In this connection, the study by Alexander, Carson, Alliger, and Cronshaw (1989) is relevant. This study, based on all available published and unpublished literature, obtained an average u_x value for GMA across a wide variety of jobs of .70, which is not far from the values in Tables 1 and 2 and is similar to the Hunter (1986) value of .67.

In summary, the evidence indicates that the artifact values used in these previously published meta-analyses are unlikely to have resulted in overestimates (or underestimates) of validity to any nontrivial degree. Hence, it is appropriate to consider the validity estimates produced by the correction for indirect range restriction more accurate than those produced by the correction for indirect range restriction. For a more general discussion of the accuracy of artifact distributions used in validity generalization studies, we refer the reader to Schmidt and Hunter (2003; pp. 54–58) and Hunter and Schmidt (2004; ch 5, pp. 169–172). For research showing that minor inaccuracies in artifact distribution values lead to only small errors in validity estimation, see Raju, Burke, Normand, and Langlois (1991). The general conclusion in this literature has been that little if any inaccuracy in validity estimates has resulted from the artifact distributions used in published meta-analyses.

With respect to theoretical conclusions about the role of personality, the major potential limitation of this study is not inherent in the study itself but in the existing research literature. If the criticisms presented earlier of the self-report personality measures that have been virtually universally used in the literature are correct, then our path analyses probably underestimate the importance of Con and/or ES in training and job performance. If so, this is not a correctable deficiency in this study. Use of more construct valid measures of Con and ES and different corrections for measurement error could lead to the conclusion that Con and ES are more important determinants of job performance than this study indicates. However, if it is the case that currently used measures of personality have severe reliability and validity deficiencies, this fact does not affect the accuracy of our estimates of operational and incremental validity in selection because these estimates apply to those personality measures. But it does mean that the validity estimates presented in this study for Con and ES may not generalize to improved measures of these personality traits that might be developed in the future.

Summary

In summary, this study indicates that when the more accurate correction for range restriction introduced by Hunter et al. (2006) is applied, validity estimates and true score correlations increase for Con, ES, and GMA. However, the increase is relatively greater for GMA, reducing the relative importance of these two personality traits in applied selection systems and in causal models of the determinants of job and training performance. Incremental validities of ES and Con are also reduced, both absolutely and relatively. These findings have important implications for understanding and theory construction and also for practical applications in employment selection. However, it is possible that future research will show that improved methods of measuring personality traits suggest a more important role for personality traits in both employment selection and theoretical models of job and training performance than is indicated in this study.

REFERENCES

Alexander RA, Carson KP, Alliger GM, Cronshaw SF. (1989). Empirical distributions of range restricted SDx values in validity studies. *Journal of Applied Psychology, 74*, 253–258.

Anderson G, Viswesvaran C. An update of the validity of personality scales in personnel selection. Unpublished manuscript.

Barrick MR, Mount MK. (1991). The Big Five personality dimensions and job performance: A meta-analysis. PERSONNEL PSYCHOLOGY, *44*, 1–26.

Barrick MR, Mount MK, Judge TA. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment, 9*, 9–30.

Barrick MR, Ryan AM. (2003). *Personality and work: Reconsidering the role of personality in organizations*. San Francisco: Jossey-Bass.

Bertua C, Anderson N, Salgado JF. (2005). The predictive validity of cognitive ability tests: A UK meta-analysis. *Journal of Occupational and Organizational Psychology, 78*, 387–409.

Borman WC, White LA, Pulakos ED, Oppler SH. (1991). Models of supervisory job performance ratings. *Journal of Applied Psychology, 76*, 863–872.

Brown KG, Le H, Schmidt FL. (2006). Specific aptitude theory revisited: Is there incremental validity for training performance? *International Journal of Selection and Assessment, 14*, 87–100.

Callender JC, Osburn HG. (1980). Development and test of a new model for validity generalization. *Journal of Applied Psychology, 65*, 543–558.

Campbell JP. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In Dunnette MD, Hough LM (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 687–732). Palo Alto, CA: Consulting Psychologists Press.

Campbell CH, Ford P, Rumsey MG, Pulakos ED, Borman WC, Felker DB, et al. (1990). Development of multiple job performance measures in a representative sample of jobs. PERSONNEL PSYCHOLOGY, *43*, 277–300.

Campbell JP, McCloy RA, Oppler SH, Sager CE. (1993). A theory of performance. In Schmitt N, Borman WC (Eds.), *Personnel selection in organization* (pp. 71–98). San Francisco: Jossey-Bass.

Campbell JP, McCloy RA, Oppler SH, Sager CE. (1993). A theory of performance. In Schmitt N, Borman WC (Eds.), *Personnel selection in organizations* (pp. 35–70). San Francisco: Jossey-Bass.

Cohen J, Cohen P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Collins JM, Schmidt FL. (1997). Can suppressor variables enhance criterion-related validity in the personality domain? *Educational and Psychological Measurement, 57*, 924–936.

Dalal RS. (2005). A meta-analysis of the relationship between organizational citizenship behavior and counterproductive work behavior. *Journal of Applied Psychology, 90*, 1241–1255.

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

Dilchert S, Ones DS, Davis RD, Rostow CD. (2007). Cognitive ability predicts objectively measured counterproductive work behaviors. *Journal of Applied Psychology, 92*, 616–627.

Dudley NM, Orvis KA, Lebiecki JE, Cortina JM. (2006). A meta-analytic investigation of conscientiousness in the prediction of job performance: Examining the intercorrelations and the incremental validity of narrow traits. *Journal of Applied Psychology*, *91*, 40–57.

Ghiselli EE. (1966). Validity of a personnel interview. PERSONNEL PSYCHOLOGY, *19*, 389–394.

Ghiselli EE. (1973). The validity of aptitude tests in personnel selection. PERSONNEL PSYCHOLOGY, *26*, 461–477.

Goldberg LR. (1990). An alternate description of personality: The Big Five factor structure. *Journal of Personality and Social Psychology, 59*, 1216–1229.

Gottfredson. (1997). Why *g* matters: The complexity of everyday life. *Intelligence, 24*, 79–132.

Hogan J, Holland B. (2003). Using theory to evaluate personality and job performance relations: A socioanalytic perspective. *Journal of Applied Psychology, 88*, 100–112.

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

Hough LM, Eaton NK, Dunnette MD, Kamp JD, McCloy RA. (1990). Criterion-related validties of personality constructs and the effect of response distortion in those validities. *Journal of Applied Psychology, 75*, 581–595.

Hough LM, Ones DS, Viswesvaran C. (1998, April). Personality correlates of managerial performance constructs. Paper presented in Page RC (Chair) *Personality Determinants of Managerial Potential Performance*, *Progression and Ascendancy*. Symposium conducted at the 13th Annual Conference of the Society for Industrial and Organizational Psychology, Dallas, TX.

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

Hunter JE. (1983). Test validation for 12,000 jobs: An application of job classification and validity generalization analysis to the General Aptitude Test Battery (GATB), Test Research Report No. 45. Washington DC: U.S. Department of Labor, Employment Service.

Hunter JE. (1985). The differential validity of ASVAB for military job performance. Report for Research Applications in partial fulfillment of Department of Defense Contract F41689-83-0025.

Hunter JE. (1986). Cognitive ability: Cognitive aptitude, job knowledge, and job performance. *Journal of Vocational Behavior, 29*, 340–362.

Hunter JE, Hunter RF. (1984). Validity and utility of alternate predictors of performance. *Psychological Bulletin, 96*, 72–98.

Hunter JE, Schmidt FL. (1990). *Methods of meta-analysis*. Newbury Park, CA: Sage.

Hunter JE, Schmidt FL. (1996). Intelligence and job performance: Economic and social implications. *Psychology, Public Policy, and Law, 2*, 447–472.

Hunter JE, Schmidt FL. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage.

Hunter JE, Schmidt FL, Le H. (2006). Implications of direct and indirect range restriction for meta-analysis methods and findings. *Journal of Applied Psychology, 91*, 594–612.

Hurtz GM, Donovan JJ. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology, 85*, 869–879.

Jensen A. (1998). *The* g *factor: The science of mental ability*. Westport, CT: Praeger.

Judge TA, Higgins C, Thoresen CJ, Barrick MR. (1999). The Big Five personality traits, general mental ability, and career success across the life span. PERSONNEL PSYCHOLOGY, *52*, 621–652.

Judge TA, Jackson CL, Shaw JC, Scott BA, Rich BL. (2007). Self-efficacy and work-related performance: The integral roles of individual differences. *Journal of Applied Psychology, 92*, 107–127.

King LM, Hunter JE, Schmidt FL. (1980). Halo in a multidimensional forced choice performance evaluation scale. *Journal of Applied Psychology, 65*, 507–516.

Le H. (2003). Correcting for indirect range restriction in meta-analysis: Testing a new meta-analytic method. Unpublished doctoral dissertation, University of Iowa, Iowa City, IA.

Le H, Schmidt FL. (2006). Correcting for indirect range restriction in meta-analysis: Testing a new analytic procedure. *Psychological Methods, 11*, 416–438.

Le H, Schmidt FL, Putka DJ. (in press). The multi-faceted nature of measurement artifacts and its implications for estimating construct level relationships. *Organizational Research Methods*.

Linn RL, Harnish DL, Dunbar SB. (1981). Corrections for range restriction: An empirical investigation of conditions resulting in conservative corrections. *Journal of Applied Psychology, 66*, 655–663.

McClelland GH, Judd CM. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin, 114*, 376–390.

McCormick EJ, DeNisi AS, Shaw JB. (1979). Use of the Position Analysis Questionnaire for establishing the job component validity of tests. *Journal of Applied Psychology, 64*, 51–56.

McCormick EJ, Jeanneret PR, Mecham RC. (1972). A study of job characteristics and job dimensions as based on the position analysis questionnaire (PAQ) [Monograph]. *Journal of Applied Psychology, 56*, 347–368.

McCrae RR, Costa PT. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 87*, 81–90.

McCrae RR, Costa PT. (1989). More reasons to adopt the five-factor model. *American Psychologist, 44*, 451–452.

McHenry JJ, Hough LM, Toquam JL, Hanson MA, Ashworth S. (1990). Project A validity results: The relationship between predictor and criterion domains. PERSONNEL PSYCHOLOGY, *43*, 335–354.

Morgeson FP, Campion MA, Dipboye RL, Hollenbeck JR, Murphy K, Schmitt N. (2007). Reconsidering the use of personality tests in personnel contexts. PERSONNEL PSYCHOLOGY, *60*, 683–729.

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

Mount MK, Barrick MR. (1995). The Big Five personality dimensions: Implications for research and practice in human resource management. *Research in Personnel and Human Resources Management, 13*, 153–200.

Mount MK, Barrick MR, Stewart GL. (1998). Five-factor model of personality and performance in jobs involving interpersonal interactions. *Human Performance*, *11*, 145–165.

Mount MK, Barrick MR. (2002). *Personal characteristics inventory™ user's manual*. Libertyville, IL: Wonderlic.

Mount MK, Barrick MR, Strauss JP. (1994). Validity of observer ratings of the Big Five personality factors. *Journal of Applied Psychology, 79*, 272–280.

Mount MK, Barrick MR, Strauss JP. (1999). The joint relationship of Conscientiousness and ability with performance: Test of the interaction hypothesis. *Journal of Management, 25*, 707–721.

Mount MK, Barrick MR, Scullen SM, Rounds J. (2005). Higher-order dimensions of the Big Five personality traits and the Big Six vocational interest types. PERSONNEL PSYCHOLOGY, *58*, 447–478.

Ones DS, Mount MK, Barrick MR, Hunter JE. (1994). Personality and job performance—a critique of the Tett, Jackson, and Rothstein (1991) meta-analysis. PERSONNEL PSYCHOLOGY, *47*, 147–156.

Ones DS, 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.

Organ DW, Ryan K. (1995). A meta-analytic review of attitudinal and dispositional predictors of organizational citizenship behavior. PERSONNEL PSYCHOLOGY, *48*, 775–802.

Pearson K. (1903). Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability and correlation of organs. *Royal Society Philosophical Transactions, Series A, 200*, 1–66.

Raju NS, Burke MJ. (1983). Two new procedures for studying validity generalization. *Journal of Applied Psychology, 68*, 382–395.

Raju NS, Burke MJ, Normand J, Langlois GM. (1991). A new meta-analysis approach. *Journal of Applied Psychology, 76*, 432–446.

Ree MJ, Carretta TR, Earles JA, Albert W. (1994). Sign changes when correcting for range restriction: A note on Pearson's and Lawley's selection formula. *Journal of Applied Psychology, 79*, 298–301.

Ree MJ, Earles JA. (1991). Predicting training success: Not much more than *g*. PERSONNEL PSYCHOLOGY, *44*, 321–332.

Ree MJ, Earles JA. (1992). Intelligence is the best predictor of job performance. *Current Directions in Psychological Science, 1*, 86–89.

Ree MJ, Earles JA, Teachout MS. (1994). Predicting job performance: Not much more than g. *Journal of Applied Psychology, 79*, 518–524.

Robertson IT, Kinder A. (1993). Personality and job competencies: The criterion-related validity of some personality variables. *Journal of Occupational and Organizational Psychology*, *66*, 225–244.

Rothstein HR. (1990). Interrater reliability of job performance ratings—growth and asymptote level with increasing opportunity to observe. *Journal of Applied Psychology*, *75*, 322–327.

Rotundo M, Sackett PR. (2002). The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: A policy-capturing approach. *Journal of Applied Psychology, 87*, 66–80.

Sackett PR, Gruys ML, Ellingson JE. (1998). Ability-personality interactions when predicting job performance. *Journal of Applied Psychology, 83*, 545–556.

Sackett PR, Ostgaard. (1994). Job-specific applicant pools and national norms for cognitive ability tests: Implications for range restrictions corrections in validation research. *Journal of Applied Psychology, 79*, 680–684.

Sackett PR, Yang H. (2000). Correction for range restriction: An expanded typology. *Journal of Applied Psychology, 85*, 112–118.

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

Salgado JF. (1998). Big five personality dimensions and job performance in Army and civil occuptations: A European persepective. *Human Performance*, *11*, 271–288.

Salgado JF. (2002). The Big Five personality dimensions and counterproductive behaviors. *International Journal of Selection and Assessment, 10*, 177–125.

Salgado JF. (2003). Predicting job performance using FFM and non-FFM personality measures. *Journal of Occupational and Organizational Psychology, 76*, 323–346.

Salgado JF, Anderson N, Moscoso S, Bertua C, de Fruyt F. (2003). International validity generalization of GMA and cognitive abilities: A European community metaanalysis. PERSONNEL PSYCHOLOGY, *56*, 573–605.

Salgado JF, Anderson N, Moscoso S, Bertua C, de Fruyt F, Rolland JP. (2003). A meta-analytic study of general mental ability validity for different occupations in the European community. *Journal of Applied Psychology, 88*, 1068–1081.

Schmidt FL. (2002). The role of general cognitive ability in job performance: Why there cannot be a debate. *Human Performance, 15*, 187–210.

Schmidt FL, Hunter JE. (1992). Development of causal models of processes determining job performance. *Current Directions in Psychological Science, 1*, 89–92.

Schmidt FL, Hunter JE. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262–274.

Schmidt FL, Hunter JE. (1999). Theory testing and measurement error. *Intelligence, 27*, 183–198.

Schmidt FL, Hunter JE. (2003). History, development, evolution, and impact of validity generalization and meta-analysis methods, 1975–2001. In Murphy KR (Ed.), *Validity generalization: A critical review* (pp. 31–66). Mahwah, NJ: Erlbaum.

Schmidt FL, Hunter JE. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology, 86*, 162–174.

Schmidt FL, Hunter JE, McKenzie RC, Muldrow TW. (1979). The impact of valid selection procedures on work-force productivity. *Journal of Applied Psychology, 64*, 609–626.

Schmidt FL, Hunter JE, Pearlman K, Hirsh HR. (1985). Forty questions about validity generalization and meta-analysis (with commentary by P.R. Sackett, M.L. Tenopyr, N. Schmitt, J. Kehoe, & S. Zedeck). PERSONNEL PSYCHOLOGY, *38*, 697–798.

Schmidt FL, Le H, Ilies R. (2003). Beyond Alpha: An empirical examination of the effects of different sources of measurement error on reliability estimates for measures of individual differences constructs. *Psychological Methods*, *8*, 206–234.

Schmidt FL, Le H, Oh I, Shaffer J. (2007). General mental ability, job performance, and red herrings: Responses to Osterman, Hauser, and Schmitt. *Academy of Management Perspectives*.

Schmidt FL, Oh I, Le H. (2006). Increasing the accuracy of corrections for range restriction: Implications for selection procedure validities and other research results. PERSONNEL PSYCHOLOGY, *59*, 281–305.

Schmidt FL, Ones DS, Hunter JE. (1992). Personnel selection. *Annual Review of Psychology, 43*, 627–670.

Schmidt FL, Viswesvaran C, Ones DS. (2000). Reliability is not validity and validity is not reliability. PERSONNEL PSYCHOLOGY, *53*, 901–912.

Schneider B. (1987). The people make the place. PERSONNEL PSYCHOLOGY, *40*, 437–453.

Schneider B. (2007). Evolution of the study and practice of personality at work. *Human Resource Management*.

Tett RP, Jackson DN, Rothstein M. (1991). Personality measures as predictors of job performance—A meta-analytic review. PERSONNEL PSYCHOLOGY, *44*, 703–742.

Thorndike RL. (1949). *Personnel selection*. New York: Wiley.

Vinchur AJ, Schippmann JS, Switzer FS, Roth PL. (1998). A meta-analytic review of predictors of job performance for salespeople. *Journal of Applied Psychology*, *83*, 586–597.

Viswesvaran C, Ones DS. (1995). Theory testing: Combining psychometric meta-analysis and structural equations modeling. PERSONNEL PSYCHOLOGY, *48*, 865–885.

Viswesvaran C, Ones DS. (2000). Measurement error in Big Five factors personality assessment: Reliability generalization across studies and measures. *Educational and Psychological Measurement, 60*, 224–235.

Viswesvaran C, Ones DS, Hough LM. (2001). Do impression management scales in personality inventories predict managerial job performance ratings? *International Journal of Selection and Assessment 9*, 277–289.

Viswesvaran C, Ones DS, Schmidt FL. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology, 81*, 557–574.

Viswesvaran C, Schmidt FL, Ones DS. (2005). Is there a general factor in ratings of job performance? A meta-analytic framework for disentangling substantive and error influences. *Journal of Applied Psychology, 90*, 108–131.

Wilk SL, Desmarais LB, Sackett PR. (1995). Gravitation to jobs commensurate with ability: Longitudinal and cross-sectional tests. *Journal of Applied Psychology, 80*, 79–85.

APPENDIX A

Explanation of the Criteria Used to Select Meta-Analyses Used in This Study

Inclusion criteria

1. The study must have been a published meta-analysis based on at least five studies or a large sample primary study.
2. The study must have reported the validity of GMA, Con, or ES for predicting overall job performance or performance in job training programs.
3. The study must have included a wide range of job types. That is, the sample could not have been limited to a specific job or a narrow range of jobs.
4. The authors must have reported the psychometric information necessary (reliabilities and range restriction ratios) necessary for reanalysis. If psychometric information was not available, meta-analytic estimates of reliabilities and range restriction ratios had to be available from another published source.

Studies included in present analysis

Barrick and Mount (1991)
Bertua et al. (2005)
Hough (1992)
Hunter (1986)
Hurtz and Donovan (2000)
Mount and Barrick (1995)
Salgado (1997)
Salgado (1998)
Salgado (2003)
Salgado, Anderson, Moscoso, Bertua, de Fruyt, and Rolland (2003)
Salgado, Anderson, Moscoso, Bertua, and de Fruyt (2003)

Excluded studies	Reason for exclusion
Hunter (1983, 1984, 1985)	These studies were based on the same basic data set as Hunter (1986). Because Hunter (1986) was the most comprehensive analysis, we chose it over the other Hunter analyses.
Anderson and Viswesvaran (1998)	This was an unpublished study that used only data from studies published between 1992 and 1997, which were included in subsequent meta-analyses.
Hough et al. (1998)	This conference presentation was later published by Viswesvaran, Ones, and Hough (2001). The independent variable was not personality but impression management scales embedded in various personality assessments.
Hough et al. (1990)	The data from Hough et al. (1990) were included in Hough (1992).
McHenry et al. (1990)	This psychometric information necessary for reanalysis was not presented in this study.
Mount et al. (1998)	This analysis included a small sample ($k = 11$) and included only jobs that required “interpersonal interaction.”
Organ and Ryan (1995)	OCB was the criterion of interest in this study. No results were presented for overall job or job training performance.
Robertson and Kinder (1993)	The criteria in this study were specific dimensions of job performance. No results were presented for overall job or job training performance.
Tett et al. (1991)	This study included several methodological errors. See Ones, Barrick, Mount, and Hunter (1994) critique for a complete critique of the methodology of this analysis.

APPENDIX A (continued)

Excluded studies	Reason for exclusion
Vinchur et al. (1998)	The sample analyzed in this study included only one job-type salespeople.
Holland and Hogan (2003)	The criteria in this study were specific dimensions of job performance. No results were presented for overall job or job training performance.

APPENDIX B
Sources of the Artifact Values Used in the Meta-Analyses

Predictor reliabilities	
Barrick and Mount (1991)	Few studies provided reliability estimates. Therefore, values were derived from test manuals (p. 10).
Bertua et al. (2005)	Derived from studies used in meta-analysis, estimates provided in inventory manuals, and from other articles that reported such reliabilities (p. 394).
Hough (1992)	No corrections were made for unreliability in the original study, nor was any artifact information presented in the paper. We used the artifact information provided by Barrick and Mount (1991) to correct these data.
Hunter (1986)	On p. 604, Hunter et al. (2006) report a mean reliability of .81 based on the USES database originally analyzed by Hunter (1983) and Hunter (1986).
Hurtz and Donovan (2000)	Derived from studies used in meta-analysis and supplemented by estimates given by inventory manuals (p. 873). Individual predictor reliability estimates were obtained via personal communication with the first author.
Mount and Barrick (1995)	No artifact information was presented in this study. We used estimates provided in Barrick and Mount (1991).
Salgado (1997)	Estimated from studies included in the meta-analysis and those provided in inventory manuals (p. 33).
Salgado (1998)	Derived from studies used in meta-analysis, estimates provided in inventory manuals, and from other articles that reported such reliabilities based on a European sample (p. 275).
Salgado (2003)	Estimated from studies included in the meta-analysis, those provided in inventory manuals, and those reported in other studies (p. 328).
Salgado, Anderson, Moscoso, Bertua, de Fruyt, and Rolland (2003)	Estimated from studies included in the meta-analysis and those provided in inventory manuals (p. 1071).
Salgado, Anderson, Moscoso, Bertua, and de Fruyt (2003)	Empirically derived from the studies in the meta-analysis ($k =$ 31) and from inventory manuals (p. 582).

APPENDIX B (continued)

Criterion reliabilities	
Barrick and Mount (1991)	Few studies provided reliability estimates. The authors used Rothstein (1990) as the basis for their criterion unreliability estimate (p. 10).
Bertua et al. (2005)	For job performance the estimate presented by Viswesvaran, Ones, and Schmidt (1996) was used. For training success, the estimate presented by Hunter and Hunter (1984) was used (p. 395).
Hough (1992)	No corrections were made for unreliability in the original study, nor was any artifact information presented in the paper. We used the artifact information provided by Barrick and Mount (1991) to correct these data.
Hunter (1986)	For job performance the estimate given by King, Hunter, and Schmidt (1980) was used (see Hunter and Hunter, 1984, p. 79). The same reliability estimate for job performance ratings was used in Hunter (1983). The reliability estimate for training performance was derived empirically from studies conducted by the U.S. Navy and was used by Ghiselli (1966, 1973; Hunter, 1986, p. 342).
Hurtz and Donovan (2000)	Combined estimates from studies included in meta-analysis with estimates provided by Rothstein (1990) (p. 873).
Mount and Barrick (1995)	No artifact information was presented in this study. We used estimates provided in Barrick and Mount (1991).
Salgado (1997)	Empirically derived from studies included in the meta analysis and from those published in *Personnel Psychology* and *Journal of Applied Psychology* (p. 33).
Salgado (1998)	This study used the criterion unreliability estimates from Salgado (1997) (p. 275).
Salgado (2003)	Used the estimate given by Viswesvaran, Ones, and Schmidt (1996) and Salgado, Anderson, Moscoso, Bertua, de Fruyt and Rolland (2003; p. 328).
Salgado, Anderson, Moscoso, Bertua, de Fruyt, and Rolland (2003)	Empirically derived from studies included in the meta-analysis. For job performance $k = 19$ (out of 69 studies); for training performance $k = 15$ (out of 60 studies).
Salgado, Anderson, Moscoso, Bertua, and de Fruyt (2003)	Empirically derived from the 93 studies included in the meta-analysis. For job performance $k = 19$; for training performance $k = 15$ (pp. 538–584).
	Range restriction ratios (u_x values)
Barrick and Mount (1991)	Empirically derived from studies included in the meta-analysis (pp. 10–11).
Bertua et al. (2005)	Empirically derived from studies included in the meta-analysis (p. 395).
Hough (1992)	No corrections were made for unreliability in the original study, nor was any artifact information presented in the paper. We used the artifact information provided by Barrick and Mount (1991) to correct these data.

APPENDIX B (continued)

	Range restriction ratios (u_x values)
Hunter (1986)	Calculated from U.S. Employment Service studies. For job performance $k = 425$; for training performance $k = 90$. These estimates are presented on page 33 of Hunter (1983) and page 343 of Hunter (1986).
Hurtz and Donovan (2000)	Used estimates given in inventory manuals (p. 873).
Mount and Barrick (1995)	No artifact information was presented in this study. We used estimates provided in Barrick and Mount (1991).
Salgado (1997)	Empirically derived from studies included in the meta-analysis and from those published in *Personnel Psychology* and *Journal of Applied Psychology* (p. 33).
Salgado (1998)	Estimated from studies included in the meta-analysis and those provided in inventory manuals (p. 275).
Salgado (2003)	Estimated from studies included in the meta-analysis and those provided in inventory manuals (p. 329).
Salgado, Anderson, Moscoso, Bertua, de Fruyt, and Rolland (2003)	Empirically derived from the studies included in the meta-analysis (p. 1072). For high complexity jobs $k = 12$ (out of 14 studies); for medium complexity $k = 13$ (of 43 studies); for low complexity $k = 5$ (of 12 studies).
Salgado, Anderson, Moscoso, Bertua, and de Fruyt (2003)	Empirically derived from the 93 studies included in the meta-analysis. For job performance $k = 20$; for training performance $k = 12$ (p. 584).

APPENDIX C

TABLE C.1
Additional Operational Validity Estimates for Direct Versus Indirect Range Restriction Corrections

Meta-analysis	(1) k	(2) N	(3) $\bar{r}$	(4) $\bar{r}_{xx_a}$	(5) $\bar{r}_{yy_i}$	(6) $\bar{u}_x$	(7) $\bar{\rho}$ DRR	(8) $\bar{\rho}$ IRR	(9) % Increase
Overall job performance criteria									
Agreeableness									
Barrick and Mount (1991)	80	11,526	.04	.76	.52	.94	.06	.06	4%
Hough (1992)	87	22,060	.05	.76	.52	.94	.07	.08	4%
Salgado (1998)	35	4,464	.01	.80	.62	.92	.01	.01	5%
Hurtz and Donovan (2000)	38	5,803	.07	.76	.59	.92	.10	.10	6%
Salgado (2003)	68	10,716	.08	.80	.52	.82	.13	.15	12%
Extraversion									
Barrick and Mount (1991)	89	12,396	.06	.76	.52	.94	.09	.09	4%
Hough (1992)	297	69,266	.07[a]	.76	.52	.94	.10	.11	4%
Salgado (1998)	39	5,116	.06	.80	.62	.92	.08	.09	5%

APPENDIX C (continued)

Meta-analysis	(1) k	(2) N	(3) $\bar{r}$	(4) $\bar{r}_{xx_a}$	(5) $\bar{r}_{yy_i}$	(6) $\bar{u}_x$	(7) $\bar{\rho}$ DRR	(8) $\bar{\rho}$ IRR	(9) % Increase
Hurtz and Donovan (2000)	37	5,809	.06	.85	.59	.92	.08	.09	3%
Salgado (2003)	75	11,940	.04	.80	.52	.86	.06	.07	9%
Openness to Experience									
Barrick and Mount (1991)	55	9,454	−.02	.76	.52	.94	−.03	−.03	4%
Hough (1992)	46	11,297	.01[b]	.76	.52	.94	.01	.02	4%
Salgado (1998)	26	4,005	.04	.80	.62	.92	.06	.06	5%
Hurtz and Donovan (2000)	33	4,881	.03	.82	.59	.92	.04	.04	4%
Salgado (2003)	48	7,562	.05	.80	.52	.85	.08	.09	10%
Training performance criteria									
Agreeableness									
Barrick and Mount (1991)	19	3,685	.06	.76	.52	.94	.09	.09	4%
Hough (1992)	7	988	.08	.76	.52	.94	.12	.12	4%
Salgado (1997)	5	415	.12	.80	.62	.92	.17	.17	4%
Extraversion									
Barrick and Mount (1991)	17	3,101	.15	.76	.52	.94	.22	.23	4%
Hough (1992)	70	8,389	.07[a]	.76	.52	.94	.10	.11	4%
Openness to Experience									
Barrick and Mount (1991)	14	2,700	.14	.76	.52	.94	.21	.21	4%
Hough (1992)	35	8,744	.02[b]	.76	.52	.94	.03	.03	4%

Note. (1) Number of validity coefficients; (2) total sample size; (3) sample size weighted mean observed validity; (4) mean predictor reliability in the unrestricted population; (5) mean criterion reliability in the restricted population; (6) mean range restriction ratio; (7) mean operational (true) validity corrected for direct range restriction; (8) mean operational (true) validity corrected for indirect range restriction; (9) percentage increase in operational validity from applying indirect range restriction correction.

[a]The observed validity used in the present analysis is the composite correlation of the observed validities of affiliation and potency presented in Hough (1992; Table 3) computed by using the meta-analytic correlation between the two (.09) from Table 2 on page 142.

[b]The observed validity used in the present analysis is the observed validity reported for intellectance presented in Hough (1992; Table 3).

TABLE C.2

Meta-Analytic Correlations Among GMA and Big Five Personality Traits

Variable		1	2	3	4	5	6
1	Con[a]	1	.33	.08	.15	.46	−.03
2	Agreeableness[a]	.39	1	.14	.22	.36	0

APPENDIX C (continued)

Variable		1	2	3	4	5	6
3	Openness to Experience[a]	.09	.17	1	.39	.17	.18
4	Extraversion[a]	.17	.26	.45	1	.21	.02
5	ES[a]	.52	.42	.19	.24	1	.07
6	GMA[b]	−.04	0	.22	.02	.09	1

Note. Correlations below the diagonal are true-score correlations, although correlations above the diagonal are observed correlations.

[a]The meta-analytic, true-score correlations among the Big Five are reported in Mount et al. (2005) and attenuated to estimate observed correlations for measurement error. Relevant psychometric information was obtained from the method section of the original paper.

[b]The meta-analytic, true-score correlations between each of the Big Five and GMA are reported in Judge et al. (2007) and attenuated to estimate observed correlations for measurement error. The psychometric information needed to compute the operational validity was not reported in the original paper, but was provided via personal communication with Timothy A. Judge.

APPENDIX D
Median Validities for All Predictors and Criteria

	Job performance		Training performance	
	DRR	IRR	DRR	IRR
GMA	.503 (.507)	.615 (.647)	.524 (.526)	.665 (.668)
Con	.201 (.210)	.209 (.222)	.250 (.235)	.260 (.245)
ES	.127 (.117)	.131 (.124)	.152 (.129)	.159 (.135)
Agreeableness[a]	.070 (.074)	.080 (.080)	.120 (.127)	.120 (.127)
Extraversion[a]	.080 (.082)	.090 (.090)	.160 (.160)	.170 (.170)
Openness to Experience[a]	.040 (.032)	.040 (.036)	.240 (.240)	.240 (.240)

Note. Mean validities are shown in parentheses.

[a]Our emphasis in this study is on operational validities of Con and ES. However, at the request of a reviewer we have included operational validities for the other Big Five factors for informational purposes.