

## **Big Five Validity and Publication Bias: Conscientiousness Worse than Assumed**

Sven Kepes  
([skepes@vcu.edu](mailto:skepes@vcu.edu))

Michael A. McDaniel  
([mamcdani@vcu.edu](mailto:mamcdani@vcu.edu))

George C. Banks  
([gcbanks@gmail.com](mailto:gcbanks@gmail.com))

Virginia Commonwealth University

Gregory M. Hurtz  
([ghurtz@csus.edu](mailto:ghurtz@csus.edu))

California State University - Sacramento

John J. Donovan  
([jdonovan@rider.edu](mailto:jdonovan@rider.edu))

Rider University

Paper presented at the annual meeting of the Society for Industrial and Organizational Psychology, Chicago, IL (April, 2011)

## **Big Five Validity and Publication Bias: Conscientiousness Worse than Assumed**

### **ABSTRACT**

The Big Five validity data from Hurtz and Donovan (2000) were analyzed to assess the presence and extent of publication bias. Evidence consistent with an inference of publication bias was found for conscientiousness, but not necessarily for the other Big Five traits.

Personality employment selection tests are popular, in part because of the general acceptance of the Big Five personality traits (e.g., Barrick & Mount, 1991; Borman et al., 2003; Hogan, 1991; Sackett & Lievens, 2008). Despite the popularity of the Big Five, the relatively weak predictive validity of personality tests remains a major concern (e.g., Hough, 1998; Hough & Oswald, 2000; Morgeson et al., 2007a, b; Murphy & Dzieweczynski, 2005). We share this concern and evaluate if the validity of the Big Five traits, although relatively low to begin with (relative to other predictors such as cognitive ability; Schmidt & Hunter, 1998), is actually overestimated due to publication bias. We use newly developed meta-analytic techniques and publication bias methods in our analyses, thereby extending prior efforts (e.g., McDaniel, Hurtz, & Donovan, 2006a). We know of no study that has examined publication bias in the validity of personality tests in this detailed and rigorous manner.

### **What is publication bias?**

Publication bias exists when studies available to reviewers are unrepresentative such that they yield systematically different results from the population of all studies (Rothstein, Sutton, & Borenstein, 2005a). Publication bias presents a great threat to the validity of meta-analytic results (Rothstein et al., 2005a), which is one of our most valuable tools for advancing evidence-based practice (Briner & Rousseau, in press; Huy, Oh, Shaffer, & Schmidt, 2007).

A typical publication bias finding is that studies with small samples and statistically insignificant results may be missing (i.e., “suppressed”) from the readily available literature (Dickerson, 2005; McDaniel, Rothstein, & Whetzel, 2006b). The editorial review process is one of the causes of study suppression and thus publication bias (e.g., rejection of studies, elimination of tangential results prior to publishing). Yet, evidence across literature areas suggests that researchers of primary studies are the principal cause of publication bias because a majority of unpublished studies are never submitted (Dickerson, 2005) or results in published studies are not fully reported (e.g., specific outcomes or sub-groups are not reported; Sutton & Pigott, 2005). Under either scenario (e.g., review process or researcher decisions), studies are not published or their results are not fully reported. Publication bias can also occur if meta-analytic researchers are unable to locate studies in the grey literature (e.g., dissertations, conference papers, government documents, organizational reports, books, etc.), which can be particularly difficult and time-consuming to locate (Hopewell, Clarke, & Mallett, 2005).

In addition, the slow dissemination of research results through published sources can manifest itself in a time-lag bias if the time to publication is shorter for studies with large samples and statistically significant effects than for studies with small samples and statistically insignificant effects (Ioannidis, 1998, 2005; Stern & Simes, 1997; Trikalinos & Ioannidis, 2005). The time-lag bias could also be due to the Proteus effect, which describes situations when studies with large effects are published earlier because they are more dramatic and interesting (Ioannidis, 2005; Trikalinos & Ioannidis, 2005). In either situation, it is proposed that earlier effect sizes (e.g., correlation coefficients between two constructs of interest) are larger than effect sizes between the same constructs obtained in later time periods (Ioannidis, 2005; Trikalinos & Ioannidis, 2005).

### **The current study**

The goal of the current study is to re-examine the Hurtz and Donovan (2000) meta-analysis of the Big Five personality traits as predictors of job performance using new meta-analytic techniques and

publication bias methods. Although publication bias could exist in all Big Five traits, we expect it to affect conscientiousness more than the other traits. First, as conscientiousness has demonstrated stronger validity than the other Big Five traits, we should be more likely to see published studies that focus primarily on the validity of conscientiousness. We suggest that such studies would be under greater pressure to demonstrate a strong validity than studies that include all Big Five traits. Thus, due to editorial-, reviewer-, and author-related reasons, studies with small magnitude correlations for conscientiousness may never get disseminated and published. Second, validity estimates for emotional stability, agreeableness, extraversion, and openness to experience are substantially smaller than for conscientiousness. Small coefficients leave little room to be explained by publication bias, even if bias is present.

## Method

### Data source

The data from Hurtz and Donovan (2000) were obtained from the authors. Although we use identical data, there are differences between the published Hurtz and Donovan (2000) study and ours. Whereas Hurtz and Donovan (2000) reported sub-group mean validities for four job families (sales, customer service, managers, and skilled and semiskilled employees), we perform separate analyses for three of these four sub-groups as the number of samples (i.e., correlations) for the managerial job family is only four, limiting the credibility of the publication bias analyses. We also perform an additional sub-group analysis for a combined sales/customer service job family. Given the content-related similarity of work between sales and customer service jobs as well as in the validity coefficients for conscientiousness (mean observed correlations: .18 and .17, estimated population means: .29 and .27, respectively), we combined these two job families for an additional analysis. Hurtz and Donovan excluded five samples ( $N=889$ ) from their sub-group analyses as they could not be classified as sales *or* customer service. Yet, it is possible to classify them as sales/customer service. This gives us an additional five samples for the combined sales/customer service job family and thus more statistical power to detect the potential presence of publication bias (the power of some methods to detect publication bias is very low with small numbers of samples; Borenstein, Hedges, Higgins, & Rothstein, 2009; Rothstein et al. 2005b).

### Analysis approach

**Meta-analytic approach and imputations.** To address statistical artifacts, we performed a psychometric meta-analysis and examined the presence of publication bias in the correlations corrected for measurement error as well as indirect range restriction (Hunter & Schmidt, 2004; Hunter, Schmidt, & Le, 2006). To calculate estimated population correlations for each observed correlation, one needs reliabilities for the personality and job performance criterion as well as information on range restriction. Our estimated population correlations reported in this study do not incorporate the correction of the coefficients for measurement error in the personality test. This is because the personality test, when used operationally, contains measurement error. However, the reliability of the personality test is needed to correct for indirect range restriction (Hunter & Schmidt, 2004; Hunter et al., 2006). The most appropriate reliability measure for supervisor ratings of job performance is inter-rater reliability (Ones, Viswesvaran, & Schmidt, 2008; Schmidt & Hunter, 1996). Therefore, we imputed the true estimated inter-rater reliability of supervisory ratings of overall job performance (.52) derived by Viswesvaran, Ones, and Schmidt (1996) when no inter-rater reliability estimate was provided.

All analyses of observed and individually corrected correlations were conducted using the CMA software (Borenstein, Hedges, Higgins, & Rothstein, 2005) because there is no software in the psychometric meta-analysis tradition that conducts all the publication bias analyses used in this study. Although the observed correlations were individually corrected for measurement error and indirect range restriction consistent with current practice in psychometric meta-analysis (Hunter & Schmidt, 2004; Hunter et al., 2006), the meta-analyses conducted in CMA differ in small ways from meta-analyses

using psychometric meta-analysis software (e.g., Schmidt & Le, 2005).<sup>1</sup> We demonstrate the small impact of these differences when comparing the mean observed correlations as reported by Hartz and Donovan (2000) with those reported in this paper. The estimated population correlations in the Hartz and Donovan results and the present results are not comparable because the original paper corrected the data for direct range restriction. The current paper uses the more appropriate indirect range restriction corrections (Hunter et al., 2006; methods for correcting for indirect range restriction were not available at the time of the original Hartz and Donovan study). Also, Hartz and Donovan used artifact distribution meta-analyses and our analyses employ individually corrected correlations (the findings of the artifact distribution approach and the individual correction of correlations result in the same conclusions; Hunter & Schmidt, 2004).

**Methods to assess publication bias.** To assess the presence of publication bias, we used five approaches, which are more suited to detect publication bias than commonly used methods (Becker, 2005; Borenstein et al., 2009; Hopewell et al., 2005; McDaniel et al., 2006b; Rothstein et al., 2005b): (a) visual examination of the funnel plot; (b) a trim and fill analysis consistent with the most established approach (Moreno et al., 2009; Sutton, 2005; Terrin, Schmid, Lau, & Olkin, 2003); (c) a cumulative meta-analysis with effect sizes sorted by precision from high to low; (d) Egger's regression interception; and (e) Begg and Mazumdar's rank correlation test. A review of these methodologies is beyond the scope of this paper (for detailed reviews see McDaniel, 2009; McDaniel et al., 2006b; Rothstein et al., 2005b). To the extent that the results are consistent across multiple publication bias methods, our conclusions are more credible.

## Results

### Conscientiousness

**Observed correlations.** Table 1 and Figures 1 and 2 display the results of our analyses. From Table 1, it can be seen that the observed mean for all job families and conscientiousness (45 samples) is .15. A trim and fill analysis of the observed correlations shows that six imputed correlations are needed in order to achieve symmetry in the funnel plot. The amount of imputed correlations indicates asymmetry in the funnel plot, indicative of publication bias (Sterne, Gavaghan, & Egger, 2005). The mean of the observed correlations combined with the six imputed correlations resulted in a mean of .12. We refer to this mean (adj.  $\bar{r}_{obs}$ ) as "the trim and fill adjusted mean observed correlation." Duval (2005) argued that a comparison of the mean observed correlation with the trim and fill adjusted mean observed correlation can help inform inferences concerning the presence of publication bias. The difference in observed means (.15 vs. .12), a difference of 20%, is relatively large and consistent with an inference of publication bias. Yet, the magnitude of the difference is only .03.

Based on the trim and fill analysis of conscientiousness for all job families combined, we conclude that there is evidence of publication bias, but the effect of this bias is relatively unimportant as a difference of .03 would probably not have much influence on decisions regarding the use of the conscientiousness measure. Begg and Mazumdar's rank correlation test ( $\tau=.00$ ,  $p=.50$ ; see Table 1) and Egger's test of the intercept (Egger's test;  $.86$ ,  $p=.13$ ) are not statistically significant and thus inconclusive (insignificant results should not be seen as evidence for the absence of publication bias; Borenstein et al., 2009; Rothstein et al., 2005b). The forest plot from the cumulative meta-analysis (available from the first

<sup>1</sup> We used CMA (Borenstein et al., 2005, 2009) to perform our analyses, partially because there is no other meta-analytic software that conducts all the publication bias analyses used in this study. CMA differs in two aspects when compared to other psychometric meta-analysis software (Schmidt & Le, 2005). First, CMA weights correlations differently. For random-effects models, CMA adjusts this weight by taking into account the variance that is not attributable to sampling error. By contrast, psychometric meta-analysis (Schmidt & Le, 2005) always uses random-effects models in which sample size is used as the study weight in a bare bones analysis (i.e., random sampling error is the only artifact considered). Due to CMA's weights for random-effects models, effects from larger sample size studies are given less relative weight when compared to psychometric meta-analysis random-effects models. Yet, the correlations between these two weights are typically high and this was true in our study. The second difference between CMA and psychometric meta-analysis is that the latter conducts analyses of correlations while CMA converts the correlations into the Fisher  $z$  metric, conducts the analyses with the Fisher  $z$ , and converts the results back into correlations. The potential difference in results between both methods is typically minimal. These differences do not affect the conclusions concerning our publication bias results.

author) shows limited indication of drift. In summary, we conclude that publication bias of the observed correlations, for data on all job families combined, shows little evidence of publication bias.

-----  
 Insert Table 1 and Figures 1 and 2 about here  
 -----

The results are similar for conscientiousness and the other job families in that there is fairly limited support for the presence of publication bias (see Table 1).

**Population correlations.** The analyses using observed correlations implicitly assume that all observed variance is solely due to sampling error and can yield incorrect conclusions when this assumption is violated (Duval, 2005; Terrin et al., 2003). Often, this assumption is doubtful, and it is likely that the Hurtz and Donovan (2000) data are heterogeneous. Three sources of heterogeneity are differences across samples in measurement error in the conscientiousness measures, measurement error in job performance, and range restriction. We did not correct for measurement error in the conscientiousness measures because an employment test has measurement error when operationally used. Yet, as discussed earlier, we did correct for measurement error in the criterion and indirect range restriction.

For conscientiousness across job families, the estimated population mean is .29 (see Table 1). This mean is higher than the population mean of .22 reported by Hurtz and Donovan (2000). This difference is primarily due to Hurtz and Donovan using corrections for direct range restriction and the current study using indirect range restriction corrections (Hunter et al., 2006). The trim and fill analysis suggests that 15 correlation coefficients are needed to be imputed in order to achieve symmetry (see Figure 1, panel a), indicative of publication bias (Sterne et al., 2005). The imputations lead to a trim and fill adjusted estimated population correlation of .16. The difference of .13 (a 45% difference) is judged substantial and is consistent with an inference of publication bias such that lower magnitude correlations are missing from the published literature. Begg and Mazumdar's rank correlation test ( $\tau=.27, p<.01$ ) and Egger's test of the intercept (1.65,  $p<.001$ ) support this conclusion.

The cumulative meta-analysis by precision also indicates the presence of publication bias (see Figure 2, panel a). The cumulative point estimate of the first five samples is .12 (cumulative  $N=2,367$ ). As less precise studies (i.e., samples with larger standard errors) are added, the cumulative estimated population mean drifts to the right side. At 51% of  $k$ , the estimate increases to .26 (cumulative  $N=5,848$ , 72% of  $N$ , cumulative  $k=23$ ). By the time all data are included, the point estimate drifts further right to .29 ( $N=8,083$ ;  $k=45$ ). The overall drift from .12 (29% of  $N$ , 11% of  $k$ ) to .29 (100% of  $N$ , 100% of  $k$ ) is quite dramatic (.17; a 58% difference). In summary, we conclude that the conscientiousness validity literature reviewed by Hurtz and Donovan has publication bias that substantially overestimates the estimated population mean. That is, the validity of conscientious is likely to be substantially lower than has been previously asserted.

Another potential source of heterogeneity in our data is moderators. Hurtz and Donovan (2000) examined job family as a moderator. We re-analyzed the job family data to evaluate the extent to which their findings were potentially distorted by publication bias. The results (see Table 1, Figures 1 and 2) are very similar, with the potential exception of the skilled and semiskilled job family. Thus, for all job families, except the skilled and semiskilled one, there appears to be publication bias in the data when using the estimated population correlations; small magnitude correlations seem to be missing from our literature.

**Time-lag bias.** Given these results, we investigated whether the time-lag bias, including the Proteus effect (Ioannidis, 1998; Sterne, Gavaghan, & Egger, 2000; Trikalinos & Ioannidis, 2005), is a potential cause for the publication bias in conscientiousness. To assess this possibility, we ran additional cumulative meta-analyses. In these analyses, we sorted the samples by year of publication (Lau & Antman, 1992; Lau, Schmid, & Chalmers, 1995). Evidence of a time-lag bias exists if the cumulative means in the forest plot drift to the negative side (i.e., left-hand side), indicating that the more recent effect sizes added to the analysis are smaller in magnitude. Consistent with the time-lag effect, the forest plots (available from the first author) for all distributions (i.e., all job families, sales/customer service,

sales, customer service, and skilled and semiskilled) show a moderate to severe left-hand drift of the cumulative point estimate (i.e., the cumulative estimated population mean). We thus conclude that the time-lag-bias, including the Proteus effect, is a likely cause for the publication bias in conscientiousness.

### **The other Big Five traits**

The results for the other Big Five personality traits show less evidence of publication bias.

### **Discussion**

This study was designed to assess whether the literature on the relation between the Big Five personality traits and job performance is affected by publication bias. Overall, our results displayed in Table 1 indicate that the observed validities for the Big Five traits are relatively small with only conscientiousness typically exceeding .10. Similarly, Hurtz and Donovan (2000) noted that the validity evidence even for conscientiousness “is not as impressive as one would expect, given the recent enthusiasm surrounding its use as a predictor of job performance” (p. 876). Our findings exacerbate this conclusion by indicating that some of Hurtz and Donovan’s (2000) observed and population mean validity estimates for the Big Five personality traits have been affected by publication bias. Of particular concern are our results regarding conscientiousness as the bias appears to be fairly severe and present for each job family; the estimated population means were overestimated by between .11 and .18 (by between 38 and 55%). Our initial argument that conscientiousness is more likely to suffer from publication bias than the other Big Five personality traits was thus supported. Additional analyses led us to the conclusion that the time-lag bias is most likely a cause of the publication bias in conscientiousness. We thus echo and extend previous cautions regarding the usage of Big Five-based selection tests (e.g., Morgeson et al., 2007a, b; Murphy & Dzieweczynski, 2005).

The differences in results for the observed correlations and population correlation estimates are likely due to differences across studies in artifacts (i.e., measurement error and range restriction; Ioannidis, 2005). Such differences distort publication bias analyses as they implicitly assume that the observed variance is due to sampling error. When using observed correlations, this assumption is likely to be violated, yielding incorrect results (Duval, 2005; Terrin et al., 2003). Thus, using the estimated population correlations (i.e., reducing the heterogeneity between samples) should increase the accuracy of publication bias methods (Duval, 2005; Rothstein et al., 2005b; Terrin et al., 2003). As corrections for indirect range restrictions have the potential to remove a significant amount of between sample heterogeneity (Hunter et al., 2006; Schmidt, Oh, & Le, 2006), it is thus not surprising that we observed differences between uncorrected (i.e., observed) and corrected (i.e., estimated population) correlations. In addition, corrections for indirect range restriction tend to lead to substantially larger population estimates (Hunter et al., 2006), reinforcing our initial explanation related to the magnitude of the validity estimates (i.e., small coefficients, coefficients that do not deviate far from zero, leave little potential to explain a possible bias, even if present). Taken together, the difference in publication bias conclusions between the observed and the estimated population correlations can, quite reasonably, be attributed to the removal of artifactual variance.

Given our results, we echo McDaniel et al.’s (2006b) call to re-examine past meta-analytic efforts for the possibility of publication bias. As we have shown, recent methodological advances allow for a rigorous assessment of publication bias that goes beyond traditionally used procedures and tests such as the problematic failsafe *N* analysis (Becker, 2005; Borenstein et al., 2009; Rothstein et al., 2005b). As meta-analytic reviews may hold the most promise in bridging the gap between research and practice by providing practitioners with a review of research findings (Briner & Rousseau, in press; Huy et al., 2007), the results and conclusions of these quantitative reviews need to be as sound and error-free as possible. Advocating the use of a specific selection test or any other management practice based on potentially erroneous results has the potential to widen the often lamented gap between research and practice (e.g., Briner & Rousseau, in press; Rousseau, 2006; Rynes, Colbert, & Brown, 2002). Thus, publication bias analysis should become a requirement in future meta-analytic efforts in the organizational sciences, just as they are in medicine (Dickerson, 2005) and certain areas of psychology (Cooper, 2003).

## References

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

- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1-23. doi: 10.1111/j.1744-6570.1991.tb00688.x
- \*Barrick, M. R., & Mount, M. K. (1993). Autonomy as a moderator of the relationships between the Big Five personality dimensions and job performance. *Journal of Applied Psychology, 78*, 111-118. doi: 10.1037/0021-9010.78.1.111
- \*Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology, 81*, 261-272. doi: 10.1037/0021-9010.81.3.261
- \*Barrick, M. R., Mount, M. K., & Strauss, J. P. (1993). Conscientiousness and performance of sales representatives: Test of the mediating effects of goal setting. *Journal of Applied Psychology, 78*, 715-722. doi: 10.1037/0021-9010.78.5.715
- \*Barrick, M. R., Mount, M. K., & Strauss, J. P. (1994). Antecedents of involuntary turnover due to a reduction in force. *Personnel Psychology, 47*, 515-535. doi: 10.1111/j.1744-6570.1994.tb01735.x
- Becker, B. J. (2005). The failsafe N or file-drawer number. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Eds.), *Publication bias in meta analysis: Prevention, assessment, and adjustments* (pp. 111-126). West Sussex, UK: Wiley.
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2005). *Comprehensive meta-analysis (Version 2)*. Englewood, NJ: Biostat.
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, West Sussex, UK: Wiley.
- Borman, W. C., Hedge, J. W., Ferstl, K. L., Kaufman, J. D., Farmer, W. L., & Bearden, R. M. (2003). Current directions and issues in personnel selection and classification. In J. J. Martocchio & G. R. Ferris (Eds.), *Research in personnel and human resources management* (Vol. 22, pp. 287-355). Oxford, UK: Elsevier Science.
- Briner, R. B., & Rousseau, D. M. (in press). Evidence-based I-O psychology: Not there yet. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 4*.
- \*Cellar, D. F., deGrange deGrendel, D. J., Klawnsky, J. D., & Miller, M. L. (1996). The validity of personality, service orientation and reading comprehension measures as predictors of flight attendant training performance. *Journal of Business and Psychology, 11*, 43-54. doi: 10.1007/BF02278254
- \*Cellar, D. F., Miller, M. L., Doverspike, D. D., & Klawnsky, J. D. (1996). Comparison of factor structures and criterion-related validity coefficients for two measures of personality based on the five factor model. *Journal of Applied Psychology, 81*, 694-704. doi: 10.1037/0021-9010.81.6.694
- Cooper, H. (2003). Editorial. *Psychological Bulletin, 129*, 3-9. doi: 10.1037/0033-2909.129.1.3
- \*Crant, J. M. (1995). The proactive personality scale and objective job performance among real estate agents. *Journal of Applied Psychology, 80*, 532-537. doi: 10.1037/0021-9010.80.4.532
- Dickerson, K. (2005). Publication bias: Recognizing the problem, understandings its origins and scope, and preventing harm. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Eds.), *Publication bias in meta analysis: Prevention, assessment, and adjustments* (pp. 11-34). West Sussex, UK: Wiley.
- Duval, S. J. (2005). The "Trim and Fill" method. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Eds.), *Publication bias in meta analysis: Prevention, assessment, and adjustments* (pp. 127-144). West Sussex, UK: Wiley.
- \*Hayes, T. L., Roehm, H. A., & Castellano, J. P. (1994). Personality correlates of success in total quality manufacturing. *Journal of Business and Psychology, 8*, 397-411. doi: 10.1007/BF02230956
- Hogan, J., Rybicki, S. L., & Borman, W. C. (1998). Relations between contextual performance, personality, and occupational advancement. *Human Performance, 11*, 189-207. doi: 10.1207/s15327043hup1102&3\_5

- \*Hogan, R. T. (1991). Personality and personality measurement. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 2, pp. 873-919). Palo Alto, CA: Consulting Psychologists Press.
- Hopewell, S., Clarke, M., & Mallett, S. (2005). Grey literature and systematic reviews. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Eds.), *Publication bias in meta analysis: Prevention, assessment, and adjustments* (pp. 48-72). West Sussex: Wiley.
- Hough, L. M. (1998). Personality at work: Issues and evidence. In M. D. Hakel (Ed.), *Beyond multiple choice: Evaluating alternatives to traditional testing for selection* (pp. 131-166). Mahwah, NJ: Lawrence Erlbaum.
- Hough, L. M., & Oswald, F. L. (2000). Personnel selection: Looking toward the future--Remembering the past. *Annual Review of Psychology*, *51*, 631-664. doi: 10.1146/annurev.psych.51.1.631
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.
- Hunter, J. E., Schmidt, F. L., & Le, H. (2006). Implications of direct and indirect range restriction for meta-analysis methods and findings. *Journal of Applied Psychology*, *91*, 594-612. doi: 10.1037/0021-9010.91.3.594
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology*, *85*, 869-879. doi: 10.1037//0021-9010.85.6.869
- Huy, L., Oh, I.-S., Shaffer, J., & Schmidt, F. L. (2007). Implications of methodological advances for the practice of personnel selection: How practitioners benefit from meta-analysis. *Academy of Management Perspectives*, *21*, 6-15.
- Ioannidis, J. P. (1998). Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *Journal of the American Medical Association*, *279*, 281-286. doi: 10.1001/jama.279.4.281
- Ioannidis, J. P. (2005). Differentiating biases from genuine heterogeneity: Distinguishing artifactual from substantive effects. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Eds.), *Publication bias in meta analysis: Prevention, assessment, and adjustments* (pp. 287-302). West Sussex, UK: Wiley.
- \*Jacobs, R. R., Conte, J. M., Day, D. V., Silva, J. M., & Harris, R. (1996). Selecting bus drivers: Multiple predictors, multiple perspectives on validity, and multiple estimates of utility. *Human Performance*, *9*, 199-217. doi: 10.1207/s15327043hup0903\_2
- \*Kolz, A. R., Cardillo, E. P., & Pena, S. A. (1998). *Cognitive ability and personality as predictors of performance and counterproductive behavior*. Paper presented at the annual meeting of the Society for Industrial and Organizational Psychology, Dallas, TX (April 1998).
- Lau, J., & Antman, E. M. (1992). Cumulative meta-analysis of therapeutic trials for myocardial infarction. *New England Journal of Medicine*, *327*, 248-254. doi: 10.1056/NEJM199207233270406
- Lau, J., Schmid, C. H., & Chalmers, T. C. (1995). Cumulative meta-analysis of clinical trials builds evidence for exemplary medical care. *Journal Of Clinical Epidemiology*, *48*, 45-57 doi: 10.1016/0895-4356(94)00106-Z
- \*Lyne, R., Sinclair, R. R., & Gerhold, C. (1997). *Personality and job performance: Matching predictor and criterion domains*. Paper presented at the annual meeting of the Society for Industrial and Organizational Psychology, St. Louis, MO (April 1997).
- \*Mabon, H. (1998). Utility aspects of personality and performance. *Human Performance*, *11*, 289-304. doi: 10.1207/s15327043hup1102&3\_9
- \*Martocchio, J. J., & Judge, T. A. (1997). Relationship between conscientiousness and learning in employee training: Mediating influences of self-deception and self-efficacy. *Journal of Applied Psychology*, *82*, 764-773. doi: 10.1037/0021-9010.82.5.764
- McDaniel, M. A. (2009). *Cumulative meta-analysis as a publication bias method*. Paper presented at the annual meeting of the Society for Industrial and Organizational Psychology, New Orleans, LA (April 2009).



- McDaniel, M. A., Hurtz, G. M., & Donovan, J. J. (2006a). An evaluation of publication bias in Big Five validity data. *Society for Industrial and Organizational Psychology, Dallas, TX (April 2006)*.
- McDaniel, M. A., Rothstein, H. R., & Whetzel, D. L. (2006b). Publication bias: A case study of four test vendors. *Personnel Psychology, 59*, 927-953. doi: 10.1111/j.1744-6570.2006.00059.x
- Moreno, S. G., Sutton, A. J., Turner, E. H., Abrams, K. R., Cooper, N. J., Palmer, T. M., et al. (2009). Novel methods to deal with publication biases: secondary analysis of antidepressant trials in the FDA trial registry database and related journal publications. *British Medical Journal (Clinical Research Ed.)*, *339*, b2981-b2981.
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007a). Are we getting fooled again? Coming to terms with limitations in the use of personality tests in personnel selection. *Personnel Psychology, 60*, 1029-1049. doi: 10.1111/j.1744-6570.2007.00100.x
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007b). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology, 60*, 683-729. doi: 10.1111/j.1744-6570.2007.00089.x
- \*Mount, M. K., Barrick, M. R., & Stewart, G. L. (1998). Five-Factor Model of personality and performance in jobs involving interpersonal interactions. *Human Performance, 11*, 145-165. doi: 10.1207/s15327043hup1102&3\_3
- \*Mount, M. K., Barrick, M. R., & Strauss, J. P. (1994). Validity of observer ratings of the big five personality factors. *Journal of Applied Psychology, 79*, 272-280. doi: 10.1037/0021-9010.79.2.272
- Murphy, K. R., & Dzieweczynski, J. L. (2005). Why don't measures of broad dimensions of personality perform better as predictors of job performance? *Human Performance, 18*, 343-357. doi: 10.1207/s15327043hup1804\_2
- Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (2008). No new terrain: Reliability and construct validity of job performance ratings. *Industrial and Organizational Psychology, 1*, 174-179. doi: 10.1111/j.1754-9434.2008.00033.x
- \*Piedmont, R. L., & Weinstein, H. P. (1994). Predicting supervisor ratings of job performance using the NEO personality inventory. *Journal of Psychology, 128*, 255. doi: 10.1080/00223980.1994.9712728
- \*Robie, C., & Ryan, A. M. (1998). *Effects of non-linearity and heteroscedasticity on the validity of conscientiousness in predicting overall job performance*. Paper presented at the annual meeting of the Society for Industrial and Organizational Psychology, Dallas, TX (April, 1998).
- \*Rose, R. M., Fogg, L. F., Helmreich, R. L., & McFadden, T. J. (1994). Psychological predictors of astronaut effectiveness. *Aviation, Space, And Environmental Medicine, 95*, 910-915.
- Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005a). Publication bias in meta-analyses. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Eds.), *Publication bias in meta analysis: Prevention, assessment, and adjustments* (pp. 1-7). West Sussex: Wiley.
- Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005b). *Publication bias in meta-analysis: Prevention, assessment, and adjustments*. West Sussex, UK: Wiley.
- Rousseau, D. M. (2006). Is there such a thing as 'evidence-based management'? *The Academy of Management Review, 31*, 256-269. doi: 10.2307/20159200
- \*Rybicki, S. L., & Klippel, D. C. (1997). *Exploring the impact of personality syndromes on job performance*. Paper presented at the annual meeting of the Society for Industrial and Organizational Psychology, St. Louis, MO (April, 1997).
- Rynes, S. L., Colbert, A. E., & Brown, K. C. (2002). HR Professionals' beliefs about effective human resource practices: correspondence between research and practice. *Human Resource Management, 41*, 149-174. doi: 10.1002/hrm.10029
- Sackett, P. R., & Lievens, F. (2008). Personnel selection. *Annual Review of Psychology, 59*, 419-450. doi: 10.1146/annurev.psych.59.103006.093716

- \*Salgado, J. F., & Rumbo, A. (1997). Personality and job performance in financial services managers. *International Journal of Selection and Assessment*, 5, 91-99. doi: 10.1111/1468-2389.00049
- Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods*, 1, 199-223. doi: 10.1037/1082-989X.1.2.199
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274. doi: 10.1037/0033-2909.124.2.262
- Schmidt, F. L., & Le, H. (2005). Hunter & Schmidt's Meta-analysis programs (Version 1.1). The University of Iowa, IA.
- Schmidt, F. L., Oh, I.-S., & 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. doi: 10.1111/j.1744-6570.2006.00037.x
- \*Schmit, M. J., Motowidlo, S. J., Degroot, T. G., Cross, T. C., & Kiker, D. S. (1996). Explaining the relationship between personality and job performance, *Society for Industrial and Organizational Psychology*. San Diego, CA (April 1996).
- Stern, J. M., & Simes, R. J. (1997). Publication bias: Evidence of delayed publication in a cohort study of clinical research projects. *British Medical Journal*, 315, 640-645.
- Sterne, J. A., Gavaghan, D., & Egger, M. (2000). Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *Journal of Clinical Epidemiology*, 53, 1119-1129. doi: 10.1016/S0895-4356(00)00242-0
- Sterne, J. A., Gavaghan, D., & Egger, M. (2005). The funnel plot. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Eds.), *Publication bias in meta analysis: Prevention, assessment, and adjustments* (pp. 75-98). West Sussex: Wiley.
- \*Stewart, G. L. (1996). Reward structure as a moderator of the relationship between extraversion and sales performance. *Journal of Applied Psychology*, 81, 619-627. doi: 10.1037/0021-9010.81.6.619
- \*Stewart, G. L. (1997). *Applicants versus incumbents: Assessing the impact of validation design on personality research*. Paper presented at the annual meeting of the Academy of Management, Boston, MA (August 1997).
- \*Stewart, G. L., & Carson, K. P. (1995). Personality dimensions and domains of service performance: A field investigation. *Journal of Business and Psychology*, 9, 365-378. doi: 10.1007/BF02230976
- \*Stewart, G. L., Carson, K. P., & Cardy, R. L. (1996). The joint effects of conscientiousness and self-leadership training on employee self-directed behavior in a service setting. *Personnel Psychology*, 49, 143-164. doi: 10.1111/j.1744-6570.1996.tb01795.x
- Sutton, A. J. (2005). Evidence concerning the consequences of publication and related biases. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Eds.), *Publication bias in meta analysis: Prevention, assessment, and adjustments* (pp. 175-192). West Sussex, UK: Wiley.
- Sutton, A. J., & Pigott, T. D. (2005). Bias in meta-analysis induced by incompletely reported studies. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Eds.), *Publication bias in meta analysis: Prevention, assessment, and adjustments* (pp. 223-240). West Sussex, UK: Wiley.
- Terrin, N., Schmid, C. H., Lau, J., & Olkin, I. (2003). Adjusting for publication bias in the presence of heterogeneity. *Statistics in Medicine*, 22, 2113-2126. doi: 10.1002/sim.1461
- Trikalinos, T. A., & Ioannidis, J. P. A. (2005). Assessing the evolution of effect sizes over time. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Eds.), *Publication bias in meta analysis: Prevention, assessment and adjustments* (pp. 241-259). West Sussex, UK: Wiley.
- \*Van Scotter, J. R., & Motowidlo, S. J. (1996). Interpersonal facilitation and job dedication as separate facets of contextual performance. *Journal of Applied Psychology*, 81, 525-531. doi: 10.1037/0021-9010.81.5.525
- Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, 81, 557-574. doi: 10.1037/0021-9010.81.5.557

**Table 1: Meta-analytically derived validity coefficients and publication bias analysis**

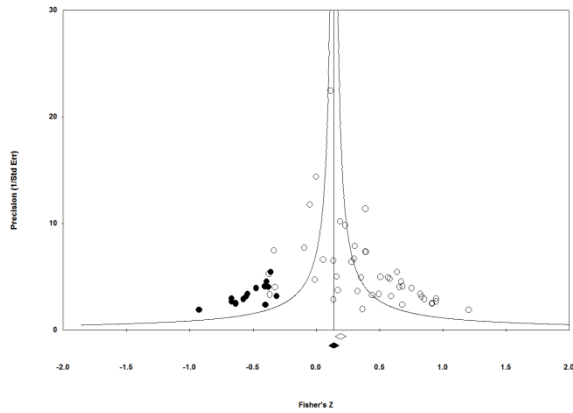
Distribution	Observed distribution								Population distribution					
	<i>N</i>	<i>k</i>	$\bar{r}_{obs}$		T&F distribution		B&M's test	Egger's test	$\hat{\rho}$	T&F distribution		B&M's test	Egger's test	
			ic	adj.	ic	adj.	$\tau$ ( <i>p</i> -value)	int ( <i>p</i> -value)		ic	adj. $\hat{\rho}$	$\tau$ ( <i>p</i> -value)	int ( <i>p</i> -value)	
<b>Conscientiousness</b>														
All job families	8,083	45	.15	(.14)	6	.12	.00 (.50)	.86 (.13)	.29	(.22)	15	.16	.27 (<.01)	1.65 (<.001)
Sales/customer service	4,107	27	.18	(n/a)	3	.16	.12 (.19)	1.10 (.17)	.36	(n/a)	10	.23	.35 (<.01)	2.18 (<.001)
Sales	1,326	10	.18	(.18)	3	.13	.11 (.33)	2.38 (.15)	.32	(.29)	5	.14	.60 (<.01)	2.74 (<.001)
Customer service	1,892	12	.18	(.17)	4	.13	.21 (.18)	3.45 (.03)	.48	(.27)	5	.30	.42 (.03)	2.67 (<.001)
Skilled and semiskilled	3,481	14	.11	(.10)	1	.09	-.05 (.39)	.20 (.44)	.20	(.17)	4	.09	.07 (.35)	1.00 (.15)
<b>Emotional Stability</b>														
All job families	5,671	37	.09	(.09)	2	.08	-.08 (.25)	.21 (.38)	.14	(.14)	4	.13	.06 (.29)	.48 (.16)
Sales/customer service	3,302	22	.11	(n/a)	1	.10	.05 (.37)	.40 (.32)	.15	(n/a)	4	.13	.13 (.19)	.97 (.07)
Sales	799	7	.09	(.09)	0	.09	-.48 (.07)	-3.10 (.13)	.13	(.15)	0	.13	.05 (.44)	-.27 (.43)
Customer service	1,614	10	.08	(.08)	3	.06	.44 (.04)	2.09 (.04)	.11	(.13)	3	.08	.47 (.03)	1.57 (.03)
Skilled and semiskilled	1,874	11	.06	(.06)	0	.06	-.24 (.16)	-.29 (.42)	.12	(.09)	1	.10	-.27 (.12)	.00 (.50)
<b>Agreeableness</b>														
All job families	6,447	40	.07	(.07)	0	.07	.02 (.42)	.08 (.47)	.10	(.13)	12	.04	.24 (.02)	1.05 (<.01)
Sales/customer service	3,567	24	.09	(n/a)	0	.09	.05 (.35)	-.43 (.35)	.14	(n/a)	10	.04	.34 (.01)	1.52 (<.01)
Sales	959	8	.04	(.03)	3	-.02	.25 (.19)	2.12 (.23)	.03	(.06)	3	-.02	.57 (.02)	4.77 (.03)
Customer service	1,719	11	.11	(.11)	1	.11	.29 (.10)	.35 (.37)	.32	(.19)	3	.29	.34 (.07)	2.67 (.13)
Skilled and semiskilled	2,385	12	.07	(.06)	0	.07	.03 (.45)	.59 (.33)	.09	(.11)	2	.05	.15 (.25)	.80 (.17)
<b>Extraversion</b>														
All job families	6,453	39	.07	(.06)	3	.05	-.01 (.46)	.63 (.21)	.10	(.10)	11	.02	.02 (.42)	1.20 (<.01)
Sales/customer service	3,573	23	.10	(n/a)	0	.10	-.09 (.26)	-.70 (.28)	.14	(n/a)	1	.14	.09 (.27)	.27 (.38)
Sales	1,044	8	.09	(.10)	0	.09	-.25 (.19)	-4.35 (.15)	.11	(.16)	0	.11	.35 (.11)	2.40 (.15)
Customer service	1,640	10	.07	(.07)	0	.07	-.04 (.43)	-.60 (.38)	.11	(.11)	0	.11	-.07 (.39)	-.31 (.41)
Skilled and semiskilled	2,385	12	.01	(.00)	0	.01	.03 (.45)	.76 (.25)	.01	(.01)	2	-.02	.15 (.25)	.79 (.10)

Distribution	Observed distribution								Population distribution							
			T&F distribution		B&M's test		Egger's test				T&F distribution		B&M's test		Egger's test	
	$N$	$k$	$\bar{r}_{obs}$	ic	adj. $\bar{r}_{obs}$	$\tau$ ( $p$ -value)	int ( $p$ -value)	$\hat{\rho}$	ic	adj. $\hat{\rho}$	$\tau$ ( $p$ -value)	int ( $p$ -value)				
Openness to Experience																
All job families	5,525	35	.04 (.04)	5	.02	.01 (.47)	.26 (.37)	.06 (.07)	8	.01	.11 (.19)	.71 (.03)				
Sales/customer service	3,156	20	.08 (n/a)	4	.06	.18 (.13)	.93 (.18)	.11 (n/a)	7	.06	.32 (.03)	1.70 (.02)				
Sales	732	6	.04 (.03)	1	.00	.40 (.13)	7.03 (.07)	.05 (.04)	3	-.07	.47 (.09)	6.40 (.01)				
Customer service	1,535	9	.10 (.10)	2	.09	.25 (.17)	1.39 (.14)	.14 (.17)	3	.11	.33 (.11)	1.50 (.05)				
Skilled and semiskilled	1,874	11	-.01 (-.01)	1	-.02	-.16 (.24)	.70 (.32)	-.01 (-.02)	1	-.02	-.05 (.41)	-.06 (.45)				

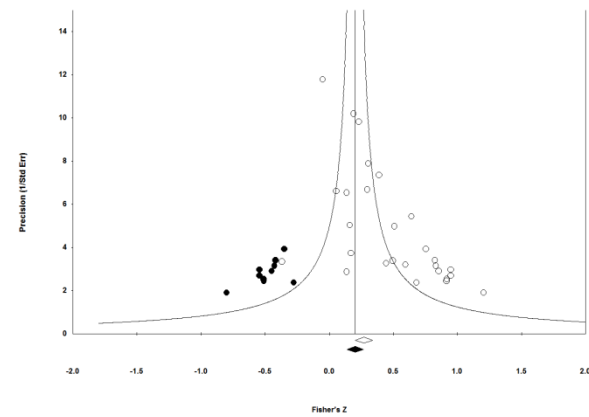
$N$  = total sample size summed across samples;  $k$  = number of samples (i.e., number of correlation coefficients);  $\bar{r}_{obs}$  = weighted mean observed correlation; ic = number of trim and fill imputed correlations; adj.  $\bar{r}_{obs}$  = trim and fill adjusted observed mean (the weighted mean of the distribution of the combined observed correlations and the imputed correlations);  $\tau$  = Kendall's tau from Begg and Mazumdar's rank correlation test; int = intercept from Egger's test of the intercept;  $p$ -value = 1-tailed  $p$ -value for Begg and Mazumdar's rank correlation test (B&M's test) and/or Egger's test of the intercept (Egger's test);  $\hat{\rho}$  = estimated population mean; adj.  $\hat{\rho}$  = trim and fill adjusted estimated population mean.

**Figure 1: Conscientiousness: Funnel plots of the estimated population correlation <sup>a</sup>**

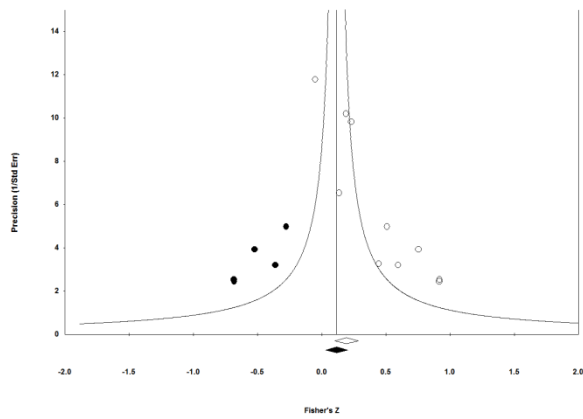
(a) All job families



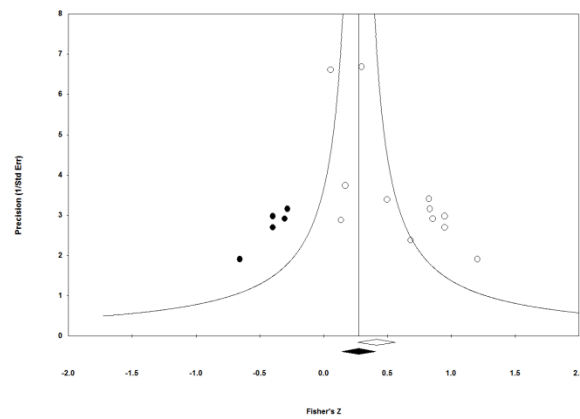
(b) Sales and customer service



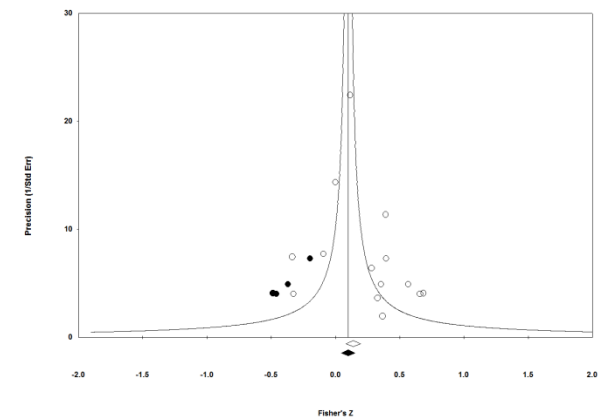
(c) Sales



(d) Customer service



(e) Skilled and semiskilled



<sup>a</sup> Imputed correlations are indicated in black.

**Figure 2: Conscientiousness: Forest plots of the cumulative meta-analyses by precision for the estimated population correlations<sup>a</sup>**

