

Publication Bias and the Validity of Conditional Reasoning Tests

George C. Banks

(gcbanks@gmail.com)

Sven Kepes

(skepes@vcu.edu)

Michael A. McDaniel

(mamcdani@vcu.edu)

Virginia Commonwealth University

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

ABSTRACT

Analyses were conducted to evaluate the possible presence and influence of publication bias on the validity of conditional reasoning tests for aggression. Multiple publication bias methods yielded results consistent with a conclusion of publication bias suggesting that the reported validity of the conditional reasoning tests may be overestimated.

The development of a new personnel selection measure is typically prompted by a desire to address limitations of current measures. Over the last decade, promising work has been conducted by James and colleagues (e.g., James, 1998; James & Mazerolle, 2002; James et al., 2005) in the development of a new personnel selection measure, known as conditional reasoning tests for aggression (CRT-A). This line of research is an attempt to develop a measure that makes it more difficult for applicants to fake (i.e., to distort responses due to the fact that the “socially desirable” answer may be apparent to test takers) and to explain construct variance not captured by conscious self-reports. Recently, a meta-analysis published by Berry, Sackett, and Tobares (2010) has offered some evidence of validity for these measures, but not as optimistic as offered earlier by James et al. (2005).

The process of accumulating validity data sufficient for a quantitative review can take years. Some research indicates that the earliest effect sizes (e.g., correlation coefficients) are larger than effect sizes obtained in later time periods (Trikalinos & Ioannidis, 2005). This phenomenon may be due to a time-lag bias, such that the time to publication is shorter for statistically-significant effects than for statistically insignificant effects (Ioannidis, 1998; Stern & Simes, 1997; Trikalinos & Ioannidis, 2005). The time-lag bias could be due to the Proteus effect (i.e., studies with large effects are published earlier because they are more interesting; Trikalinos & Ioannidis, 2005). Thus, validity studies in relatively new literatures may be subject to a temporal bias, such that initial findings overestimate the validity of a test, and resulting in publication bias. More formally, publication bias exists to the extent that research findings available to consumers of research are unrepresentative of all research results (McDaniel, Rothstein, & Whetzel, 2006; Rothstein, Sutton, & Borenstein, 2005b). The typical result of publication bias is an overestimate of effect sizes (Dickersin, 2005; McDaniel et al., 2006). It has been proposed that publication bias is one of the greatest threats to the validity of meta-analysis (Rothstein et al., 2005b), which is one of our most important tools for advancing science and evidence-based management (Briner & Rousseau, in press). Therefore, confidence in the validity and robustness of our meta-analytic results is contingent upon the extent to which publication bias influences our research.

Additional causes of publication bias

One of the more frequently discussed causes of publication bias are editor and reviewer decisions (Dickersin, 2005). In this case, a study is rejected by the editorial review process because the sample size was small, the results were not statistically significant (Greenwald, 1975), the results were contrary to theory, contrary to the position of the editor/reviewers, or contrary to purportedly well-established knowledge (Davis, 1971; Dickersin, 2005). However, the problem can also occur because of author decisions. For instance, an author may never submit a study to a journal or a conference because the study had a small sample size, the results were statistically insignificant, contrary to theory, contrary to trends of past research, or contrary to the position of the author.

Organizational constraints are another potential cause of publication bias. For example, results may not be submitted because the data are proprietary in nature or organizational researchers with monetary interests in a product (e.g., a commercially-marketed employment test) may suppress a study because the results could damage sales of the product (McDaniel et al., 2006). A final cause of publication bias concerns accessing grey literature, such as conference papers, dissertations, and technical reports. In this instance, it is possible that a paper is not identified in a systematic literature search.

Conditional reasoning tests of aggression (CRT-A)

CRT-A scales present test takers with a problem that needs to be solved using reason and judgment. CRT-A scales ask test takers to select a response that most logically flows from the problem presented. When taking these tests, participants believe that their critical intellectual skills are evaluated as they are asked to select a logical conclusion (James et al., 2005). However, CRT-A scales assess precursors to aggression (James et al., 2005); they identify systematic biases in what test takers believe are rational analyses (James & Mazerolle, 2002). James et al. (2005) identified several systematic biases, including the hostile attribution bias, potency bias, and retribution bias among others. Examples of the rationalizations for the systematic biases include a justifiable correction of an injustice, self-defense in response to being attacked, an attempt to restore self-respect, and a legitimate strike against oppression. These rationalizations serve as defense mechanisms for hostile test takers to protect a favorable sense of the self (Baumeister, Campbell, Krueger, & Vohs, 2003; Baumeister, Dale, & Sommer, 1998; Westen, 1998). On the other hand, non-hostile test takers tend to make use of rationalizations that illustrate prosocial ideologies and tendencies (James et al., 2005).

The current study

The current study presents an important contribution to scholarship in light of the possible effects of publication bias in general, and specifically time-lag bias, including the Proteus effect, in the CRT-A data. A second important reason to conduct the analysis is that the *American Psychological Association* style manual (APA, 2010) recommends an assessment of publication bias in all meta-analytic reviews. Whereas the Berry et al. (2010) article did not provide such an analysis, the current paper compliments Berry et al.'s meta-analysis by filling this gap in their analysis. Thus, we conduct the analyses using the same data set as Berry et al. (2010).¹ A third important contribution of this article is to guide evidence-based practice. Based on the quantitative summaries by James et al. (2005) and Berry et al. (2010), the CRT-A scales are likely to be attractive to practitioners, particularly because their findings suggest that the scales have a higher validity than personality tests and it may be more difficult for test takers to distort responses in a socially desirable fashion. However, should the publication bias analyses suggest that the CRT-A validities are substantially lower than believed, CRT-A scales may become less attractive to practitioners. Therefore, this study serves as a sensitivity analysis to the findings presented by James et al. (2005) and Berry et al. (2010).

Method

Data source

The data used in this study were provided by Berry et al. (2010).

Meta-analysis procedure

There are many methods for evaluating publication bias. A detailed review of the advantages and limitations of each method is beyond the scope of this paper (for reviews see McDaniel et al., 2006; Rothstein, Sutton, & Borenstein, 2005a). Some methods have clearly established themselves as superior to others. For example, the trim and fill analysis (Duval & Tweedie, 2000a, 2000b) has demonstrated clear advantages over the failsafe N (Becker, 2005; McDaniel et al., 2006). The use of cumulative meta-analysis has also been offered as a useful tool for the detection of publication bias (Borenstein, Hedges, Higgins, & Rothstein, 2009; McDaniel, 2009). Several methods for evaluating publication bias are likely useful in identifying

the presence and influence of publication bias. To the extent that multiple methods are consistent with an inference of publication bias, one can have greater confidence in the conclusion of publication bias.

Multiple publication bias methods were used in this study. *Comprehensive Meta-analysis* (CMA; Borenstein, Hedges, Higgins, & Rothstein, 2005) was used to complete the following analyses:² (1) a meta-analysis of observed correlations using both random-effects and fixed-effects models; (2) a trim and fill analysis on random-effects and fixed-effects models; (3) a cumulative meta-analysis with correlations sorted by precision from high to low (i.e., the inverse of the standard error of the correlation) and a cumulative meta-analysis with correlations sorted by publication date from most distant to most recent; (4) Egger's regression intercept (Using precision to predict the standardized effect size; standardized effect size is effect size divided by the standard error); (5) Begg and Mazumdar's rank correlation test; and finally, (6) a one-study removed sensitivity analysis.

Results

Table 1 displays the studies included in our analyses. There were 21 unique samples with a total of 3,820 individuals. The first two columns display the study ID and the author(s) of the study. The next three columns display the sample size for each study, the observed correlation, and the criterion type (e.g., CWB or job performance).

Table 2 displays the summary of the meta-analysis of the observed correlations as well as the publication bias analyses and tests (i.e., trim and fill analysis, Begg and Mazumdar's rank correlation test, and Egger's test of the intercept) by criterion. The table contains results for both fixed- and random-effects models to demonstrate the minor differences between both techniques (the conclusions of the analyses do not change). Yet, we focus on the results from the random-effects models, and discuss their implications as such models provide more accurate estimates given our data and research inquiry (Borenstein et al., 2009; Hunter & Schmidt, 2004).

The results displayed in Table 2 are consistent with an inference of publication bias. For instance, in the random-effects analysis using CWB as the sole criterion, seven correlations were imputed to create a symmetrical distribution (Figure 1). The mean observed correlation (.22) was adjusted downward (.08), resulting in a difference of .14. Similarly, the observed 95% confidence interval (.12 to .31) was adjusted downward (-.02 to .19) and includes zero. The difference in observed means ($\Delta \bar{r}_{obs}$) is judged as substantial (.14; a difference of 64%), which is consistent with an inference of severe publication bias and the conclusion that small magnitude correlations are likely to be missing from our journals and available grey literature. Therefore, the data are consistent with an inference that publication bias is likely present and that the best estimate of the observed mean is likely to be substantially less than .22. Egger's test supports this conclusion (4.83, $p < .00$), but Begg and Mazumdar's test does not (this should not be interpreted as evidence for the absence of publication bias; Borenstein et al., 2005; Borenstein et al., 2009)

Cumulative meta-analysis. Figure 2 displays the forest plots of our cumulative meta-analyses of the CWB samples; the cumulative meta-analyses by precision (panel (a)) and by publication year (panel (b)). In both instances, the respective meta-analysis is re-computed as samples are added one by one. For the cumulative meta-analysis by precision, the most precise samples (i.e., the largest samples) are added first, the least precise samples (i.e., the smallest samples) last. The forest plot for the CWB samples (see Figure 2, panel (a)) indicates a clear right-hand drift. The cumulative point estimated of the first two samples is -.01 (cumulative N

$[N_{cum}] = 1,119$; 35% of N). As less precise samples are added, the cumulative point estimate drifts to the right-hand side such that by the time the cumulative sample size reaches 53% ($N_{cum} = 1,888$), the point estimate is .24. After that, the point estimate stabilizes at around .22, which is also the final estimate. The overall drift from -.01 (35% of N) to .22 (100% of N) is dramatic ($\Delta = .23$). Consistent with our conclusion based on the trim and fill analysis, the cumulative meta-analysis by precision suggests that the publicly available literature on CRT-A scales likely has publication bias. A very similar but potentially more severe pattern emerges for the distribution of all the samples.

To test for the presence of a time-lag effect, we sorted the samples by the year they were published and conducted a second cumulative meta-analysis by year of publication. Consistent with the time-lag effect, including the Proteus effect, the forest plot (Figure 2, panel (b)) shows a clear left-hand side drift, suggesting that samples with large effects are published earlier, potentially because they are more interesting, or because small effect-size samples take substantially longer to appear in print (Ioannidis, 1998; Stern & Simes, 1997; Trikalinos & Ioannidis, 2005). In conclusion, several separate analyses presented in this study are in agreement with an inference of publication bias.

Sub-group analyses. Two sub-group analyses were conducted using the CWB samples. First, published studies (e.g., studies obtained from journals and test manuals) showed an observed correlation of .29, compared to .17 from the grey literature. Second, a trim and fill analysis was conducted separately on student samples and non-student samples. The results indicated the presence of publication bias in the 11 student samples. Five studies were imputed to make the distribution of samples symmetrical, resulting in a difference of .04 between the observed and adjusted correlation (the adjusted confidence interval to include zero). The distribution of the six non-student samples required the imputation of three studies to make the distribution of samples symmetrical and that resulted in a change of .13 from the observed to the adjusted correlation, and it adjusted the confidence interval to include zero.

One study removed. We conducted a one-study removed analysis to evaluate the sensitivity of our results. The results are virtually identical for all incidents of the analysis. In only one out of 17 incidents did our results change substantially, when we removed sample 10 ($N = 770$; Walton, 2004). However, the CRT-A scale used in sample 10 is identical to most of the other scales, the observed correlation (-.06) is not an apparent outlier (i.e., several studies reported negative correlations, two with correlations at or below -.10), etc. The only difference is its size ($N = 770$), which is substantially larger than the other samples. This gives the sample more weight in our analyses because it is the most precise estimate of the population parameter (Borenstein et al., 2009; Hunter & Schmidt, 2004) making us reluctant to advocate its deletion. Finally, as this study is an extension of Berry et al.'s (2010) meta-analysis, which included this sample, we included it too.

Discussion

This study assessed whether publication bias has affected the validity estimates of the CRT-A scales. Overall, our results are consistent with an inference of publication bias such that the results of Berry et al. (2010) overestimate the validity of the CRT-A scales. For the distribution of the CWB samples, the trim and fill analysis estimated an adjusted observed mean (.08) of less than half of the unadjusted mean observed correlation (.22). In addition, the observed confidence interval (.12 to .31) was adjusted down (-.02 to .19) and included zero. The difference in means (.14) is consistent with an inference of publication bias, such that the results

identified by James et al. (2005) and Berry et al. (2010) are likely to overestimate the CRT-A validities. Furthermore, the differences could be judged severe in that practitioners might decide to use the test if the validity was .22 for CWB, but few might decide to use the test if the validity was .08.

The cumulative meta-analyses provided additional evidence consistent with an inference of publication bias. When sorting by precision, the forest plot showed a clear right-hand side drift, suggesting that small sample, small magnitude studies are missing from our available literature. In addition, when sorting by publication date, a temporal pattern in the parameter estimate emerged indicative of the time-lag bias (e.g., Trikalinos & Ioannidis, 2005).

Limitations

There are two primary limitations to this study. First, there is a possibility that unknown heterogeneity resulted in inaccuracies in the trim and fill analysis. However, other publication bias methods (i.e., cumulative meta-analysis as well as Egger's test of the intercept) provided evidence consistent with a conclusion of publication bias, supporting the trim and fill results. Second, there are only a limited number of publicly available samples using the CRT-A scales, partly because research on CRT-A scales is relatively recent. Therefore, the conclusions reached in this study are restricted by the number of studies available. Future research will be needed to provide greater confidence in the validity of CRT-A scales.

Conclusion

The use of CRT-A scales for personnel selection is relatively recent in comparison to more established tests (e.g., cognitive ability, personality, or situational judgment tests). Our results using multiple methods to detect and assess publication bias are consistent with an inference of publication bias, such that the James et al. (2005) review and the Berry et al. (2010) meta-analysis have likely overestimated the validity of the CRT-A scales. We recommend that our results be replicated as more data accumulate and more unpublished studies are located. We support the recommendations of others (e.g., APA, 2010; McDaniel et al., 2006; Rothstein et al., 2005b) to address the issue of publication bias in all meta-analyses, and to report any findings in spite of the result. The confidence we have in the validity and robustness of meta-analytic results is reliant upon the extent to which publication bias influences our research. Finally, we call for additional research using the CRT-A scales.

References

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

- APA. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: American Psychological Association.
- Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does high self esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Journal of the American Psychological Society, 4*, 2-44. doi: 10.1111/1529-1006.01431.
- Baumeister, R. F., Dale, K., & Sommer, K. L. (1998). Freudian defense mechanisms and empirical findings in modern social psychology: Reaction formation, projection, displacement, undoing isolation, sublimation, and denial. *Journal of Personality, 66*, 1082-1124. doi: 10.1111/1467-6494.00043.
- 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.
- Berry, C. M., Sackett, P. R., & Tobares, V. (2010). A meta-analysis of conditional reasoning tests of aggression. *Personnel Psychology, 63*, 361-384. doi: 10.1111/j.1744-6570.2010.01173.x.
- *Bing, M. N., Stewart, S. M., Davison, H. K., Green, P. D., McIntyre, M. D., & James, L. R. (2007). An integrative typology of personality assessment for aggression: Implications for predicting counterproductive workplace behavior. *Journal of Applied Psychology, 92*, 722-744. doi: 10.1037/0021-9010.92.4.722.
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive Meta-analysis version 2*. Englewood, N.J.: Biostat.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex: John Wiley & Sons, Ltd.
- 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*(1).
- Davis, M. S. (1971). That's interesting! Towards a phenomenology of sociology and a sociology of phenomenology. *Philosophy of the Social Sciences, 1*, 309-344. doi: 10.1177/004839317100100211
- Dickersin, K. (2005). Recognizing the problem, understanding its origins and scope, and preventing harm. In H. R. Rothstein & A. J. Sutton (Eds.), *Publication bias in meta analysis: Prevention, assessment and adjustments* (pp. 11-34). West Sussex, UK: Wiley.
- Duval, S. J., & Tweedie, R. L. (2000a). A non-parametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*, 89-98.
- Duval, S. J., & Tweedie, R. L. (2000b). Trim and fill: A simple funnel plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*, 276-284. doi: 10.1111/j.0006-1341X.2000.00455.x.
- *Frost, B. C. (2002). *Implicit and explicit personality: An integrative approach to predicting aggressive behavior in a field setting*. (Doctoral Dissertation, University of Tennessee).
- *Frost, B. C. (2005). *Implicit and explicit personality: An integrative approach to predicting aggressive behavior in a field setting*. (Doctoral Dissertation, University of Tennessee). ProQuest Dissertations, AAT 3188875.
- Greenwald, A. G. (1975). Consequences of prejudice against the null hypothesis. *Psychological Bulletin, 82*, 1-20.

- *Hawes, S. R. (2000). *A comparison of biodata, ability, and a conditional reasoning test as predictors of reliable behavior in the workplace*. (Doctoral Dissertation, University of Tennessee). ProQuest Dissertations, AAT 9996356.
- *Hornick, C. W., Fox, K. A., Axton, T. R., & Wyatt, B. S. (1999). *The relative contribution of conditional reasoning and multiple intelligence measures in predicting firefighter and law enforcement officer job performance*. Unpublished manuscript.
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.
- 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.
- James, L. R. (1998). Measurement of personality via conditional reasoning. *Organizational Research Methods*, 1, 131-163. doi: 110.1177/109442819812001.
- James, L. R., & Mazerolle, M. D. (2002). *Personality at work*. Thousand Oaks, CA: Sage.
- *James, L. R., & McIntyre, M. D. (2000). *Conditional reasoning test of aggression test manual*. Knoxville, TN: Innovative assessment technology, LLC.
- James, L. R., McIntyre, M. D., Glisson, C. A., Green, P. D., Patton, T. W., LeBreton, J. M., et al. (2005). A conditional reasoning measure for aggression. *Organizational Research Methods*, 8, 69-99. doi: 10.1177/1094428104272182.
- *LeBreton, J. M. (2002). *Use of differential framing to measure implicit social cognitions associated with aggression*. (Doctoral dissertation, University of Tennessee). ProQuest Dissertations, AAT 3054126.
- McDaniel, M. A. (2009). Cumulative meta-analysis as a publication bias method. *Paper presented at the Paper presented at the 24th annual meeting of the Society for Industrial and Organizational Psychology, New Orleans*.
- McDaniel, M. A., Rothstein, H. R., & Whetzel, D. L. (2006). Publication bias: A case study of four test vendors. *Personnel Psychology*, 59, 927-953. doi: .
- *Patton, T. W. (1998). *Measuring personal reliability via conditional reasoning: Identifying people who will work reliably*. (Doctoral Dissertation, University of Tennessee). ProQuest Dissertations, AAT 9936290.
- Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005a). *Publication bias as a threat to validity*. West Sussex, UK: Wiley.
- Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005b). Publication bias in meta-analysis. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment, and adjustments* (pp. 1-7). West Sussex, UK: Wiley.
- *Russell, S. M., James, L. R., & McIntyre, M. D. (2004, April). *Predicting lying, cheating, and defiance in an internet-based environment*. Poster presented at the 19th Annual Conference for the Society for Industrial and Organizational Psychology, Chicago, IL.
- *Sablinski, C. J., & Mitchell, T. R. (2006, May). Post hoc analyses of interactions between the aggression questionnaire and conditional reasoning test of aggression. In Lebreton JM (Chair), *Bringing the implicit personality into I-O psychology*. Symposium presented at the 21st Annual Conference for the Society for Industrial and Organizational Psychology, Dallas, TX.
- 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.

- Trikalinos, T. A., & Ioannidis, J. P. (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.
- *Walton, W. R. (2004). *Justification of antisocial behavior*. (Doctoral Dissertation, University of Tennessee). ProQuest Dissertations, AAT 3148313.
- Westen, D. (1998). The scientific legacy of Sigmund Freud: Toward a psychodynamically informed psychological science *Psychological Bulletin*, *124*, 333-371. doi: 310.1037/0033-2909.1124.1033.1333.

Table 1: Summary of samples included in the meta-analysis

Study ID	Author(s)	Sample size	Observed correlation	Criterion
1.	James & McIntyre (2000)	105	.34	CWB
2.	James & McIntyre (2000)	111	.43	CWB
3.	Hawes (2000)	349	.06	CWB
4.	Bing et al. (2007)	176	.07	CWB
5.	Patton (1998)	100	.38	CWB
6.	James & McIntyre (2000)	135	.31	CWB
7.	Russell et al. (2004)	191	.40	CWB
8.	Frost (2005)	183	.40	CWB
9.	Frost (2002)	191	.25	CWB
10.	Walton (2004)	770	-.06	CWB
11.	Lebreton (2002)	105	-.11	CWB
12.	Lebreton (2002)	121	-.10	CWB
13.	Lebreton (2002)	130	.13	CWB
14.	Sablynski & Mitchell (2006)	95	.28	CWB
15.	James & McIntyre (2000)	188	.37	CWB
16.	Bing et al. (2007)	225	.22	CWB
17.	Bing et al. (2007)	62	.31	CWB
18.	Hawes (2000)	395	.12	Job performance
19.	Hornick et al. (1999)	68	-.03	Job performance
20.	Hornick et al. (1999)	52	.31	Job performance
21.	Hornick et al. (1999)	68	.31	Job performance

Table 2: Meta-analytically derived validity coefficients and publication bias analysis^a

Criterion	Meta-analysis results					Trim and fill results ^c					Additional tests			
	Fixed-effects		Random-effects			Fixed-effects			Random-effects		Egger	B&M		
	k	\bar{r}_{obs}	95% CI	\bar{r}_{obs}	95% CI	ic	adj. \bar{r}_{obs}	95% CI	$\Delta \bar{r}_{obs}$	adj. \bar{r}_{obs}			95% CI	$\Delta \bar{r}_{obs}$
Overall perf. ^b	21	.16	.13 to .19	.21	.13 to .29	8	.08	.05 to .10	.08	.09	.00 to .18	.12	3.66**	0.04
CWB	17	.16	.13 to .19	.22	.12 to .31	7	.06	.03 to .10	.10	.08	-.02 to .19	.14	4.83**	0.07
Job perf.	4	.14	.06 to .22	.16	.02 to .30	1	.12	.04 to .20	.02	.13	-.01 to .26	.03	1.12	0.00

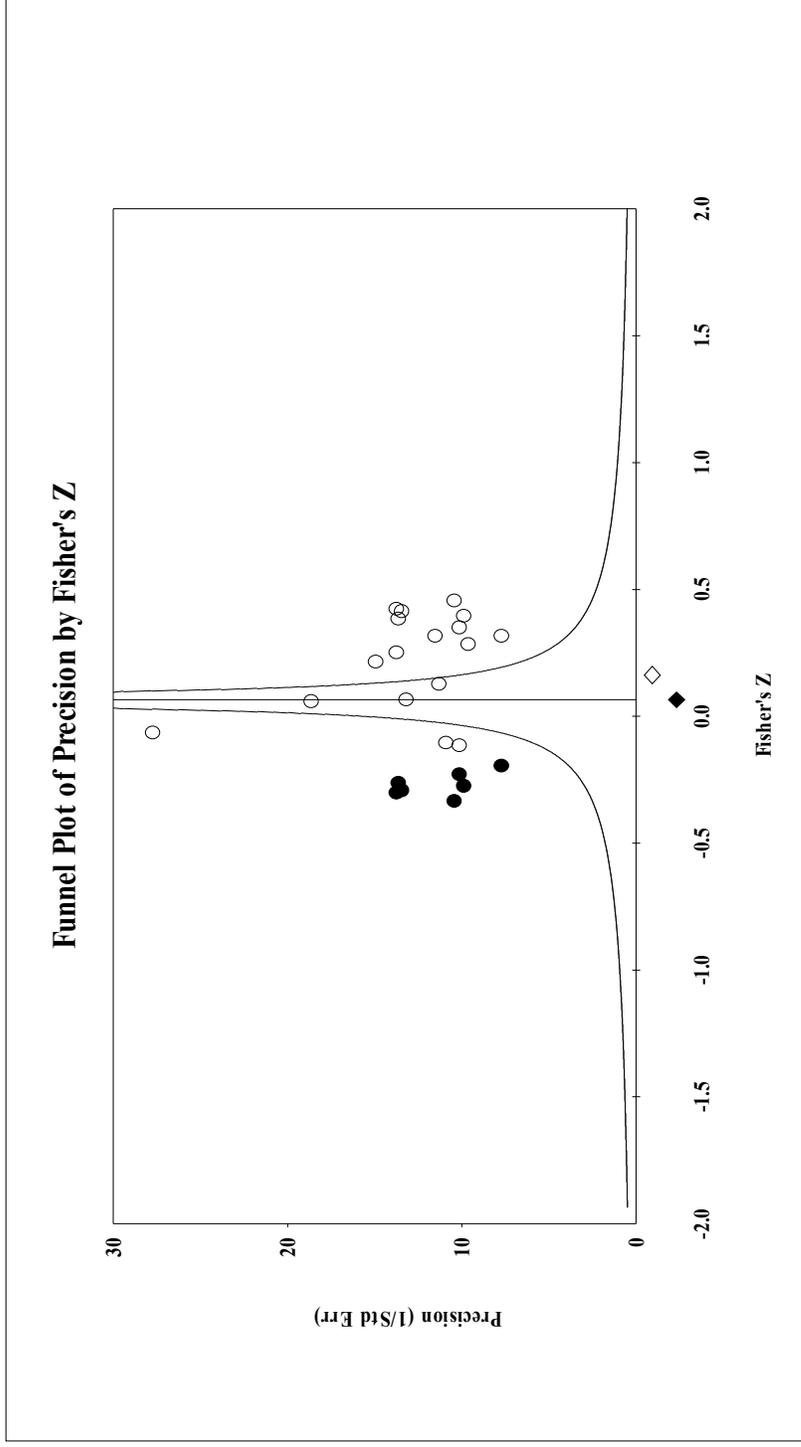
^a k = number of samples (i.e., number of correlation coefficients); \bar{r}_{obs} = weighted mean observed correlation; 95% CI = 95% confidence interval; ic = number of trim and fill imputed correlations; adj. \bar{r}_{obs} = trim and fill adjusted observed mean (the weighted mean of the distribution of the combined observed correlations and the imputed correlations); adj. 95% CI = trim and fill adjusted 95% confidence interval; $\Delta \bar{r}_{obs}$ = difference between \bar{r}_{obs} and adj. \bar{r}_{obs} ; Egger = Egger's test of the intercept (Using precision to predict the standardized effect size); B&M = Begg and Mazumdar's rank correlation between effect size and precision (Kendall's tau b).

** = significant at .00 (1-tailed).

^b Overall perf. = CWB and job perf. (i.e., entire sample)

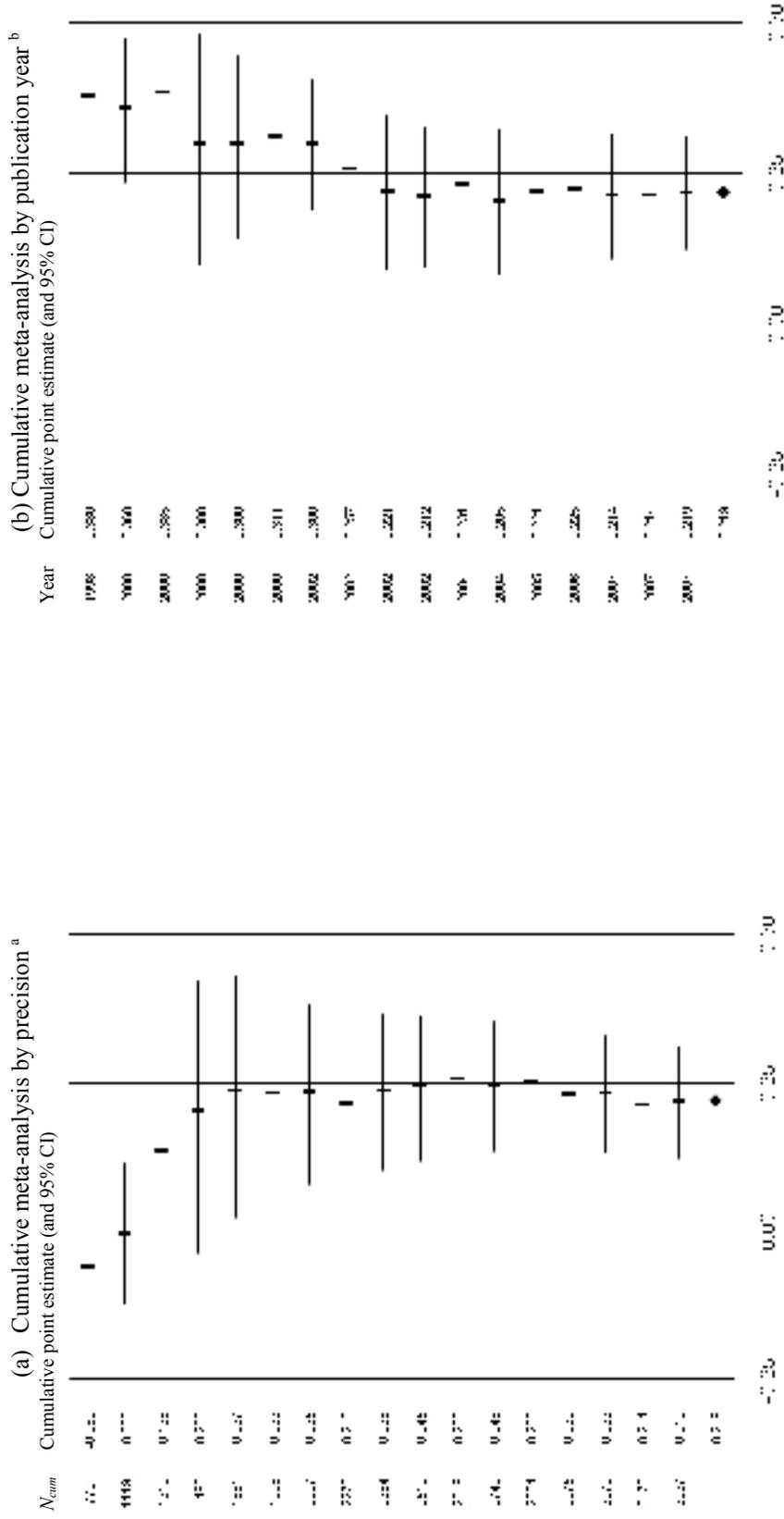
^c The number of studies imputed is the same for both the fixed-effects and random-effects models because a fixed-effects model is the optimal technique to trim and fill studies from the funnel plot in the iterative process (Cooper & Hedges, 1994; Duval, 2005).

Figure 1: Funnel plot^a



^a Funnel plot of CWB observed correlations ($k=17$) with trim and fill imputation. Imputed correlations are indicated in black ($ic=7$).

Figure 2: Forest plots of the cumulative meta-analysis for counterproductive work behaviors



^a The sample sizes and the point estimate in each row are cumulative (i.e., N_{cum} =Cumulative sample size)

^b The point estimate in each row is cumulative; Year=publication year

PUBLICATION BIAS

¹ We thank Christopher Berry for providing the data in an automated form. We note that the omission of the APA-recommended publication bias analyses in the Berry et al. article might be due to the article being accepted prior to the release of the 2010 APA style manual.

² CMA software implements meta-analysis consistent with the traditions of non-psychometric meta-analyses (Borenstein et al., 2009). For analysis of fixed-effects models, CMA uses weights derived from sample size. Specifically, it uses precision (the inverse of the sampling error variance). In its analysis of random effects models, CMA modifies this weight to also reflect the variance that is not attributable to sampling error. Psychometric meta-analysis (e.g., Schmidt & Le, 2005), is always a random effects model but uses sample size as the study weight in its analysis of observed correlations. In this study, the correlation between sample size and the precision weight in the CMA fixed-model analysis was .95. For the CMA random effects analysis, the correlation between sample size and the random-effects weights was .77. As a result, the findings of CMA's fixed-effects model are virtually identical to the results of a random-effects psychometric meta-analysis (both analyses yield a mean correlation of .16 for CWB criteria). The CMA random effects model yields a mean correlation of .21 for CWB criteria. These differences do not affect the conclusions concerning publication bias.