A Trim and Fill Examination of the Extent of Publication Bias in Communication Research

Christopher J. Carpenter

Western Illinois University

Cj-carpenter2@wiu.edu

Twitter: @DrCJCarpenter

This is an Accepted Manuscript of an article published by Taylor & Francis in Communication
Methods and Measures on March 2, 2012, available online:

<https://doi.org/10.1080/19312458.2011.651347>

Abstract

Publication bias can occur when a study is not published in an academic journal because the study did not find a statistically significant result. It can bias meta-analytic estimates upwards because the meta-analyses are missing studies that provide small estimates of the effect size under study. This paper describes the trim and fill technique of estimating the impact of publication bias on meta-analyses (Duval and Tweedie, 2000a, 2000b) and provides a worked example. A sample of 68 meta-analyses of communication research were tested for publication bias using the trim and fill technique. The results suggest that most communication meta-analyses are not substantially affected by publication bias.

A Trim and Fill Examination of the Extent of Publication Bias in Communication Research

According to Sutton (2005), publication bias is “a tendency toward preparation, submission, and publication of research findings based on the nature and direction of the research results” (p. 13). Researchers generally conduct more studies than they publish. Research conducted on social scientists’ publication choices found that when researchers fail to detect a statistically significant effect, they are less likely to attempt to publish that study (Cooper, DeNeve, & Charlton, 1997). Dissertations and conference papers can be obtained to help fill the gap but the number of studies that scholars produce and never publish or present at a conference is unknowable. Even when scholars attempt to publish these findings, many journal editors and reviewers are less likely to publish null findings. When null findings occur, there is a presumption that either the relationship in the population is negligible and thus of little interest or that the researcher conducted a flawed study. Unfortunately, some editors may forget that null findings are inevitable due to sampling error and thus fail to recognize the importance of publishing all the findings on a given statistical relationship. As a result, the scientific community does not learn about these null findings.

The problem of publication bias can be traced back centuries. Dickersin (2005) described scientists worrying about the failure to publish “successful” studies as far back as the 17th century. Researchers have discovered publication bias in clinical trials of medical treatments (Sutton, Duval, Tweedie, Abrams, & Jones, 2000), sociology (Gerber & Malhotra, 2008), criminology (Rothstein, 2008), and experimental psychology (Cooper et al., 1997). The problem is not limited to quantitative research. An examination of qualitative studies found the proportion of qualitative papers presented at conferences that are later published is as low as the proportion of quantitative papers that are later published (Petticrew, et al., 2008). Recently, Levine, Asada,

and Carpenter (2009) found evidence in meta-analyses of communication research consistent with publication bias.

Meta-analyses estimate the relationship between variables by calculating a sample size weighted effect size across all known studies of that relationship. Although publication bias affects any attempt to summarize a research literature, its impact on meta-analysis is particularly acute because meta-analyses seek to provide a more precise estimate of a statistical relationship than narrative reviews. When there are unpublished studies of a relationship between two variables that find very small effects, those studies are unlikely to be included in a meta-analysis of that relationship. The result is an upward bias in the meta-analytic estimate of the true effect size (Sutton, 2005). Attempts to estimate the impact of missing studies published in the medical sciences have only just begun diffusing in meta-analyses of clinical trials (Parekh-Burke et al., 2011). The rate of adoption in communication has been even slower as very few meta-analyses in communication research have attempted to address this problem (Levine, et al., 2009).

Critics have questioned the value of social science research due to publication bias (Delgado-Romero & Howard, 2005). Ioannidis (2007) argued that “most published research findings are false” (p. 696) because the widespread reliance on statistical significance testing causes many non-statistically significant studies to remain unpublished. Even popular press outlets are calling attention to the problem of publication bias due to well-publicized cases of clinical trials that were unpublished and resulted in the approval of unsafe new pharmaceuticals (Lehrer, 2010). Lehrer questions the ability of researchers to provide the public with accurate information. If communication researchers wish to justify the value of their work, increased attention to publication bias is necessary.

One method of detecting and estimating the impact of publication bias due to failure to publish low effect size studies with small sample sizes is the trim and fill procedure (Duval and Tweedie, 2000a, 2000b). This technique is a relatively easy to implement algorithm that first estimates how many unpublished studies there are concerning a given statistical relationship. Then the procedure estimates what the meta-analytic effect size estimate for that relationship would be if those studies had been included. The purpose of this paper is twofold: to introduce this technique to communication researchers and to use it to estimate how inflated the effect size estimates are in a sample of communication meta-analyses.

Trim and Fill Procedure

The trim and fill procedure is based on an early attempt to graphically identify publication bias in meta-analyses called the funnel plot (Light and Pillemer, 1984). Sampling error theory predicts that when a researcher draws a small sample, the obtained effect size is more likely to vary from the true population effect size than if that researcher had drawn a larger sample. Theoretically, when a meta-analyst collects a large number of studies on the relationship between two variables, she or he can expect to find more variation in effect sizes amongst small-sample studies than amongst large-sample studies. If one plots the observed effect size on the x-axis and sample size or standard error on the y-axis, one should see an upside down funnel shape. If there is publication bias, it will appear as though a piece of the funnel around the smaller effect sizes on the left side of the funnel will be missing. If the average effect size is negative, the missing piece would be on the right side of the funnel. Negative effect sizes are found when an inverse relationship is detected using a correlation or when Cohen's d is used to indicate that the control group scored higher than the experimental group. Figure 1 shows a funnel plot for the first meta-analysis reported by Rafaeli-Mor & Steinberg (2002). Examination

of that funnel plot shows that the distribution is asymmetrical, suggesting publication bias.

Unfortunately, simple visual examination of the funnel plot as a test of publication bias has been shown to be error-prone (Terrin, Schmid, & Lau, 2005). The funnel plot is therefore limited as a reliable method of estimating publication bias.

The trim and fill procedure uses the logic of the funnel plot to estimate how many studies it would take for the funnel to be symmetrical when there are studies with small samples and effect sizes missing from the left side of the funnel plot (Duval and Tweedie, 2000a, 2000b). The trim and fill procedure is implemented in a series of computationally simple steps. Duvall (2005) describes the steps in an easy to follow system and they are listed and illustrated in Table 1 using the data from the first Rafaeli-Mor & Steinberg (2002) meta-analysis. First, the weighted mean effect size is calculated and used as an estimate of the population effect size. Next, that estimate of the population effect size is subtracted from each effect size in the meta-analysis. The absolute values of these obtained values are then calculated while noting which were initially negative. These absolute values are then ranked from smallest to largest. The ranks of all of the values that were positive before the absolute values were taken are then summed. This method is based on the Wilcoxon statistic (1945) that was designed to determine if pairs of ranked scores were different to a statistically significant degree such that one score tended to be higher than the other. In this case, however, the estimate of the population mean was subtracted from each study effect size rather than subtracting a second score from the same subject as per the original Wilcoxon method. The algorithm is attempting to determine if there are an unequal number of positive and negative differences between the pairs of scores. The greater this statistic, the more asymmetrical the distribution around the mean effect size is.

When using the original Wilcoxon statistic, the total is compared to a table to determine statistical significance, but in this case the trim and fill is not attempting to determine statistical significance. Instead, the sum is used to estimate the number of missing studies that were suppressed due to publication bias. This sum (S_{rank}) is entered into the Duval and Tweedie (2000b) formula along with the total number of studies (n) to calculate the (L_0) estimator¹:

$$L_0 = [4S_{\text{rank}} - n(n+1)]/[2n - 1]$$

The formula produces the L_0 estimate of the number of studies missing from the meta-analysis due to publication bias. L_0 is then rounded to the nearest whole number to provide an estimate of the number of studies that caused asymmetry in the funnel plot. This rounded estimate is referred to as k_0 . The list of ranked values that was used previously to calculate S_{rank} is then examined and k_0 of the largest positive values that represent the largest effect sizes in the meta-analysis are removed from the list of effect sizes in the meta-analysis. Then, using the reduced list of effect sizes, the weighted mean effect size is again estimated without those effect sizes included.

The preceding steps are then repeated with the remaining effect sizes using the new estimate of the population effect size. After sufficient iterations have occurred for k_0 to reach zero, the trimming portion is complete. Duval (2005) reports that it usually only takes two or three iterations before $k_0 = 0$ in realistic data sets. All of the trimmed effect sizes are then put back into the sample of studies in the meta-analysis. Next, for every effect size that was trimmed, a new effect size that is just as discrepant from the new estimate of the population effect size is added in on the left side of the distribution. These new effect sizes are assigned the same sample sizes as each of the effect sizes they were mirrored from. Figure 1 shows a funnel plot for the Rafaeli-Mor & Steinberg (2002) data with the missing studies added in as white circles on the left side of the distribution. The weighted mean estimate of the population effect size and the

new weighted mean variance estimates are calculated from this new data set with the hypothetically missing studies included.

The trim and fill procedure provides two pieces of useful information to the meta-analyst. First, the number of studies that may be missing from the obtained corpus of studies can be estimated. Second, the possible effect on the estimate of the population effect size can be produced. Borenstein, Hedges, Higgins, and Rothstein (2009) note that there are generally three possible outcomes of this estimate: the estimate of the effect size is not changed at all, the estimate is changed trivially (as is the case in the example in Table 1), or the effect size estimate is affected so greatly it changes the interpretation of the relationship under study. If the trim and fill procedure suggests a substantial difference between the trim and fill estimate and the original estimate, the meta-analysis should be interpreted cautiously.

Alternate Methods of Detecting Publication Bias

One of the earliest methods of detecting publication bias is the failsafe N proposed by Rosenthal (1979). Levine et al. (2009) found that although tests for publication bias in communication meta-analyses were rare, when they were conducted, the failsafe N was the most common. The formula proposed by Rosenthal seeks to estimate how many studies finding a null result would have to be added to the body of effect sizes being examined before the estimate would no longer be statistically significant. Presumably if the number of studies required to reduce the effect size estimate to nonsignificance is large, then the estimate is unlikely to be substantially biased by a few unpublished studies.

There are several limitations to the failsafe N that reduces its utility for detecting publication bias. Becker (2005) outlined several of these limitations. First, it assumes that all of the unpublished studies found effects of exactly zero, a highly unlikely outcome. Some

unpublished studies might have found effects in the opposite direction and still others may have simply found small effects in the predicted direction that did not reach statistical significance. Additionally, the failsafe N does not indicate how large of an impact publication bias may have had on the effect size estimate. Work examining the ability of the failsafe N to detect unpublished studies found that it produces inaccurate estimates (Pham, Platt, McAuley, Klassen, & Moher, 2001). Becker concludes that given the myriad alternatives that have been developed since, the failsafe N should be abandoned.

Other techniques use regression to determine the extent of funnel plot asymmetry by testing for a linear relationship between estimates of each study's precision (the inverse of its standard error) on each study's effect size (Sterne & Egger, 2005). One of the most well-known is Egger's Linear Regression method (Egger, Davey Smith, Schneider, & Minder, 1997). This method regresses the standard normal deviation (the standardized effect size divided by its standard error) onto the inverse of its standard error. Egger et al. argue that in the absence of publication bias the intercept of the regression line will pass through zero. A statistically significant negative intercept suggests publication bias.

This technique is relatively easy to implement, but it is not without its limitations. One obvious limitation is that it can detect publication bias but it cannot determine the impact of publication bias on the effect size estimate. Additionally, research on the accuracy of the Egger's test has found high Type 1 error rates such that publication bias is detected where none exists (Sutton, 2009). Type 1 error rates are stronger when the effect size is dichotomous or when the number of studies is small. Macaskil, Walter, and Irwig (2001) found unacceptably high Type 1 error rates and recommended that if meta-analysts use the Egger regression test that they supplement it with the trim and fill procedure. Although it may be helpful to use a variety of

methods to detect publication bias and compare them in any given meta-analysis, the trim and fill procedure provides the most information.

Study Rationale

The trim and fill procedure has been used extensively since its inception and is now the most popular method of assessing the impact of publication bias in meta-analysis (Sutton, 2009). Yet, communication meta-analysts have rarely employed the trim and fill procedure. A search of communication journals turned up only two uses of the trim and fill procedure (Carpenter & Boster, 2009; Sun, Pan, & Shen, 2008). Therefore, with the exception of these two articles, the amount of publication bias in meta-analytic estimates in communication meta-analyses is largely unknown. Levine et al. (2009) found that sample sizes and effect sizes are often negatively correlated in communication meta-analyses. Because such a relationship indicates an asymmetrical funnel plot, they argue that many communication meta-analyses are overestimating population effect size due to publication bias. Yet, it remains unknown if the effect size estimates found in those meta-analyses are overestimated by a trivial or substantial magnitude.

The current study conducted a trim and fill analysis of the same set of communication meta-analyses examined by Levine et al. (2009). The Levine et al. finding that publication bias exists in many communication meta-analyses was predicted to be conceptually replicated using the trim and fill procedure estimates. In order to test that hypothesis, trim and fill analyses were conducted to estimate the number of missing studies and the impact on the estimated population effect size for each meta-analysis in the sample. The Levine et al. method of examining the $n-r$ correlation was also calculated to compare the estimates from the $n-r$ correlation to the estimates from the trim and fill procedure.

Method

Sample

The same sample of meta-analysis articles that Levine et al. (2009) sampled was used for this analysis in order to provide a conceptual replication of their work. They collected 51 articles containing meta-analyses on communication phenomena. The meta-analyses they gathered all contained tables showing the effect sizes and sample sizes of the articles that were meta-analyzed. Additionally, only meta-analyses with ten or more studies were included here because the Cochran Handbook for Systematic Reviews of Interventions (Higgins & Green, 2011) argues that tests of funnel plot asymmetry are underpowered when there are less than ten studies. This sample included 49 articles. From these articles, sample effect size estimates and sample size data were available for 68 independent meta-analyses. For a reference list of these articles see Levine et al. (2009)².

Coding

All of the independent effect sizes and their sample sizes were recorded from each of the 68 meta-analyses. In most cases, the meta-analyses reported correlation coefficients. When Cohen's d s were reported, they were converted to correlations.

Analysis

The metafor software package (Viechtbauer, 2010) was used to conduct the trim and fill analyses and to create the funnel plots displayed in Figure 1³. The L_0 estimator of the number of missing studies was used for the trim and fill analyses (Duval and Tweedie, 2000a, 2000b). The Hunter and Schmidt (2004) method of estimating the population effect size was employed to

estimate the average effect sizes for each meta-analyses because Anker, Reinhart, and Feeley (2010) found that it possessed the strongest validity of the meta-analysis methods they examined. The correlation between sample size and the effect size (n-r correlation) in each meta-analysis was calculated using the method described by Levine et al. (2009). The absolute value of each correlation was taken and a correlation between these values and their associated sample sizes was calculated.

Results

The incidence of publication bias from the trim and fill analysis will be reported first. Then the relationship between the results of the Levine et al. (2009) n-r correlation and the trim and fill procedure will be examined. Table 2 displays the results of both analyses for each meta-analysis.

First, of the 67 meta-analyses examined, the trim and fill procedure estimated that 19 were missing studies due to publication bias. Among those 19 meta-analyses, the trim and fill procedure estimated an average of 8.58 ($SD = 8.78$) studies were missing from each due to publication bias. The average decrease in the effect size among those 19 studies was $r = .05$. The decrease in effect sizes ranged from .01 to .15: 10 were below .05, seven were between .05 and .1, and two were above .1. Overall, the trim and fill procedure estimated that 27% of the obtained meta-analyses were inflated due to publication bias but very few of those were inflated enough to have substantially altered the conclusions of the meta-analysis in question.

Levine et al. (2009) claim that as the negative relationship between sample size and effect size grows more substantial, the more likely publication bias is present. Therefore, the trim and fill procedure estimates were predicted to be correlated with the n-r correlation. The relationship between the n-r correlation for each meta-analysis and the number of missing studies the trim

and fill procedure estimated was $r = -.25, p < .05$. The relationship between the n-r correlation and the reduction in the size of the correlation estimated by the trim and fill procedure was $r = -.35, p < .05$. This finding is consistent with a moderate level of agreement between the two estimators. These results suggest that as the negative n-r correlation grows more substantial, there are likely to be more missing studies and a greater inflation of the effect size due to publication bias.

Discussion

This study of publication bias in 67 communication meta-analyses produced several important findings. Initially, although there may be a few cases of publication bias substantially inflating effect sizes in communication meta-analyses, in most cases the impact of publication bias is minimal or non-existent. This finding is inconsistent with a trim and fill analysis of medical clinical trials meta-analyses (Sutton et al. 2000). That study found that 54% of the clinical trials had evidence of publication bias compared to the 27% found here. It appears that publication bias is a greater problem in medical research than in communication research.

The relative rarity with which publication bias substantially inflated the meta-analytic estimates in communication research is consistent with Hunter and Schmidt's (2004) claim that publication bias should be rare in meta-analyses of social scientific research. They argue that in most social science studies, researchers test multiple hypotheses. As a result, even though the statistical power to detect medium sized statistically significant effects is small in most social scientific studies, a study that tests multiple hypotheses has more chances to confirm at least one of those hypotheses. Even when sampling error causes several Type 2 errors in a single study, at least one or two of the hypotheses is still likely to be accurately confirmed. Although there may be a bias against studies that fail to reject the null hypothesis, the statistically significant results

detected for some of the hypotheses in a given study might be enough for the article to be published. Additionally, Kerr (1998) argues that many social scientists change their hypotheses to match the data they obtained (although Kerr argues this is an unethical practice). If many social scientists change their hypotheses to fit the statistically significant relationships they find, publication bias may be less rampant. These two tendencies may reduce publication bias in the social sciences.

The two meta-analyses that the trim and fill procedure found were overestimated by more than .1 bear closer examination. The original estimate for the Priess and Allen (1998) meta-analysis of the effect of incentives on counterattitudinal advocacy is $r = -.06$ such that adding incentives caused counterattitudinal advocacy to be less potent. The trim and fill estimate of the effect of incentives on counterattitudinal advocacy is $r = .09$, a reversal of the direction of the effect. A similar finding occurred for the Allen, Emmers-Sommer, and Crowell (2002) meta-analysis of the relationship between talking about condoms prior to sex and the use of condoms. The initial estimate of that relationship was small and positive, $r = .05$ whereas the trim and fill estimate was small and negative, $r = -.06$. On the other hand, both articles noted that the effects were heterogeneous (there were undetected moderators that cause the population effect size to vary). The trim and fill method may sometimes detect publication bias where none exists if the effect is heterogeneous (Peters, Sutton, Jones, Abrams, & Rushton, 2007). Additionally, both meta-analyses included some unpublished work such as master's theses, doctoral dissertations, and conference papers. There may have been more unpublished work missing, but the authors did clearly make an effort to locate those missing studies.

Finally, the results of the trim and fill procedure, though not completely congruent with the n-r correlation method of detecting publication bias (Levine et al., 2009), were moderately

correlated with the n-r results. Levine et al. argue that their results support a substantial amount of publication bias whereas the trim and fill analysis found substantially less. This discrepancy may be due to the trim and fill procedure producing more precise estimates of the effects of publication bias whereas the n-r correlation lacks specific criteria to determine if publication bias is present. Given that the trim and fill procedure does contain criteria for the identification of publication bias, it is to be preferred to the n-r correlation as a strategy for detecting bias.

Limitations

This analysis and the trim and fill procedure itself faces several limitations. Ideally a random sample of communication meta-analyses would have been obtained and analyzed for this project. The sample used here may be considered a convenience sample of only those communication meta-analyses that contained tables with the sample sizes and effect sizes for the included studies. The results of this analysis may not generalize to all communication meta-analyses if this sample is different from the rest of the population of communication meta-analyses. The limitation to articles that reported all of the effect sizes and sample sizes may have limited the sample to higher quality meta-analyses that conducted more thorough searches for unpublished articles. On the other hand, this sample did cover a wide variety of topics, sample sizes, and effect size estimates (see Levine et al., 2009 for more details concerning the sample).

The trim and fill procedure itself has several important limitations. The simulation studies conducted by Duval and Tweedie (2000a, 2000b) found the trim and fill procedure was effective when the set of studies was homogeneous (i.e. the meta-analysis produced an estimate of a single population effect size that was not moderated by other variables). Other simulation studies have found that when the set of studies are heterogeneous, the trim and fill procedure may report publication bias where it does not exist (Peters et al., 2007). Many meta-analyses of

communication research find such heterogeneous estimates. Due to these limitations, the trim and fill procedure is recommended as a sensitivity test rather than a statistical correction. On the other hand, even in the presence of heterogeneity, if the trim and fill procedure indicates little or no publication bias, then the estimate is likely to be accurate. One solution is to separate the studies by moderator and only employ the trim and fill procedure on the subsets of studies that are homogeneous. This remedy, however, will only increase accuracy if the initial meta-analysis is very large and the subsets contain enough studies to produce a meaningful estimate of publication bias (Peters, Sutton, Jones, & Abrams, 2010).

Additionally, the trim and fill procedure, like the funnel plot that inspired it, is inaccurate if there is a relationship between sample size and study quality due to some systematic difference between the studies (Sterne, Becker, & Egger, 2005). Small studies may find stronger effects because rigorous methods are sometimes time-intensive and necessitate smaller sample sizes due to scarce resources. Alternatively large studies may be conducted using external funding that would allow the collection of larger data sets with rigorous methods. Researchers conducting large studies may also simply conduct their research more carefully due to the large amount of resources committed to those studies. Sutton (2009) argues that either or both of these possibilities may easily occur in a sample of studies. Meta-analysts should be aware of any systematic differences in study quality that might covary with sample sizes in the sample of studies they are examining. If such systematic covariance is identified, the trim and fill procedure and other methods based on asymmetry in funnel plots may not be appropriate (Rothstein, 2008).

Finally, it is important to remember that publication bias is not the only form of bias that can reduce the accuracy of effect size estimates. Hunter and Schmidt (2004) describe many artifacts of studies that can bias estimates including measurement error, range restriction, weak

construct validity, artificial dichotomization, and others. They provide corrections for many of these artifacts. The trim and fill, however, should not be considered a correction.

Conclusion

Although this analysis did not find evidence consistent with publication bias for most of the meta-analyses under review, this article should not be interpreted as evidence that publication bias should not be a concern of communication meta-analysts. The trim and fill procedure did detect substantial publication bias in several cases that would alter the interpretation of those meta-analyses. Future meta-analysts would be advised to examine the $n-r$ correlation as Levine et al. (2009) recommend and employ the trim and fill procedure. Furthermore, researchers conducting meta-analyses should keep up with the growing literature on publication bias as more valid algorithms than the trim and fill procedure may be developed in the future. The best solution is for meta-analysts to make every effort to find unpublished articles by searching for dissertations, conference papers, technical reports, and other reports of research that were unpublished. Journal editors could help by publishing well executed studies that report null results. If communication researchers wish to use meta-analyses to help answer questions of theoretical and societal importance, they must take care that they are not producing inaccurately high estimates due to publication bias.

References

- Allen, M., D'Alessio, D., Emmers, T. M., & Gebhardt, L. (1996). The role of educational briefings in mitigating effects of experimental exposure to violent sexually explicit material: A meta-analysis. *The Journal of Sex Research*, 33, 135-141.
- Allen, M., Emmers-Sommer, T. M., & Crowell, T. L. (2002). Couples negotiating safer sex behaviors: A meta-analysis of the impact of conversation and gender. In M. Allen, R. W. Preiss, B. M. Gayle, & N. A. Burrell (Eds.), *Interpersonal communication research: Advances through meta-analysis* (pp. 263-279). Mahwah, NJ: Lawrence Erlbaum Associates.
- Anker, A., Reinhart, A., & Feeley, T. (2010). Meta-analysis of meta-analyses in communication: Comparing fixed effects and random effects analysis models. *Communication Quarterly*, 58, 257-278.
- Bax, L. (2011). MIX 2.0. Professional software for meta-analysis in Excel. Version 2.0.1.4. BiostatXL [Software]. Available from <http://www.meta-analysis-made-easy.com>.
- Becker, B. J. (2005). Failsafe *N* or file-drawer number. In H. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta analysis: Prevention, assessment and adjustments* (pp. 111-125). Chichester, UK: Wiley.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2005). Comprehensive Meta-analysis Version 2 [Software]. Available from <http://www.meta-analysis.com>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.
- Carpenter, C. J., & Boster, F. J. (2009). A meta-analysis of the effectiveness of the disrupt-then-reframe compliance gaining technique. *Communication Reports*, 22, 55-62.

- Cooper, H., DeNeve, K., Charlton, K. (1997). Finding the missing science: The fate of studies submitted for review by a human subjects committee. *Psychological Methods*, 2, 447-452.
- Cruz, M. G. (1998). Explicit and implicit conclusions in persuasive messages. In M. Allen & R. W. Preiss (Eds.), *Persuasion: Advances through meta-analysis* (pp. 217-230). Cresskill, NJ: Hampton Press.
- Delgado-Romero, E. A., & Howard, G. S. (2005). Finding and correcting flawed research literatures. *The Humanistic Psychologist*, 33, 293-303.
- Dickersin, K. (2005). Publication bias: Recognizing the problem, understanding its origins and scope, and preventing harm. In H. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta analysis: Prevention, assessment and adjustments* (pp. 12-33). Chichester, UK: Wiley.
- Duval, S. (2005). The trim and fill method. In H. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication Bias in Meta Analysis: Prevention, Assessment and Adjustments* (pp. 127-144). Chichester, UK: Wiley.
- Duval, S., & 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., & 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.
- Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315, 629-634.
- Gerber, A. S., & Malhotra, N. (2008). Publication bias in empirical sociological research: Do

- arbitrary significance levels distort published results? *Sociological Methods & Research*, 47, 3-30.
- Higgins, J. P. T., & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*. Retrieved from www.cochrane-handbook.org.
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting for error and bias in research findings*. Thousand Oaks, CA: Sage Publications.
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Med*, 2, e124.
- Jennions, M. D., & Møller, A. P. (2002). Publication bias in ecology and evolution: An empirical assessment using the 'trim and fill' method. *Biological Reviews of the Cambridge Philosophical Society*, 77, 211-222.
- Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, 2, 196-217.
- Lehrer, J. (2010, December). The truth wears off: Is there something wrong with the scientific method? *New Yorker*, 52.
- Levine, T. R., Asada, K. J., & Carpenter, C. (2009). Sample sizes and effect sizes are negatively correlated in meta-analysis: Evidence and implications of a publication bias against nonsignificant findings. *Communication Monographs*, 76, 286-302.
- Light, R. & Pillemer, D. (1984). *Summing up: The science of reviewing research*. Cambridge, MA: Harvard University Press.
- Macaskill, P., Walter, S. D., & Irwig, L. (2001). A comparison of methods to detect publication bias in meta-analysis. *Statistics in Medicine*, 20, 641-654.
- Parekh-Bhurke, S. et al. (2011). Uptake of methods to deal with publication bias in systematic

- reviews has increased over time, but there is still much scope for improvement. *Journal of Clinical Epidemiology*, 64, 349-357.
- Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2007). Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. *Statistics in Medicine*, 26, 4544-4562.
- Peters, J. L., Sutton, A. J., Jones, D. R., & Abrams, K. R. (2010). Assessing publication bias in meta-analyses in the presence of between-study heterogeneity. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 173, 575-591.
- Petticrew, M., Egan, M., Thomson, H., Hamilton, V., Kunkler, R., & Roberts, H. (2008). Publication bias in qualitative research: What becomes of qualitative research presented at conferences? *Journal of Epidemiology and Community Health*, 62, 552-554.
- Pham, B., Platt, R., McAuley, L., Klassen, T. P., & Moher, D. (2001). Is there a “best” way to detect and minimize publication bias?: An empirical evaluation. *Evaluation & the Health Professions*, 24, 109-125.
- Preiss, R. W., & Allen, M. (1998). Performing counterattitudinal advocacy: The persuasive impact of incentives. In M. Allen & R. W. Preiss (Eds.), *Persuasion: Advances through meta-analysis*(pp. 241-242). Cresskill, NJ: Hampton Press.
- Rafaeli-Mor, E., & Steinberg, J. (2002). Self-complexity and well-being: A review and research synthesis. *Personality and Social Psychology Review*, 6, 31-58.
- Rosenthal, R. (1979). The ‘file drawer problem’ and tolerance for null results. *Psychological Bulletin*, 86, 638-641.
- Rothstein, H. R. (2008). Publication bias as a threat to the validity of meta-analytic results. *Journal of Experimental Criminology*, 4, 61-81.

- Sterne, J. A. C., Becker, B. J., & Egger, M. (2005). The funnel plot. In H. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication Bias in Meta Analysis: Prevention, Assessment and Adjustments* (pp. 75-98). Chichester, UK: Wiley.
- Sterne, J. A. C. & Egger, M. (2005). Regression methods to detect publication and other bias in meta-analysis. In H. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication Bias in Meta Analysis: Prevention, Assessment and Adjustments* (pp. 99-110). Chichester, UK: Wiley.
- Sun, Y., Pan, Z., & Shen, L. (2008). Understanding the third-person perception: Evidence from a meta-analysis. *Journal of Communication*, 58, 280-300.
- Sutton, A. J. (2005). Evidence concerning the consequences of publication bias and related biases. In H. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication Bias in Meta Analysis: Prevention, Assessment and Adjustments* (pp. 175-192). Chichester, UK: Wiley.
- Sutton, A. J. (2009). Publication bias. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed.). New York, NY: Sage.
- Sutton, A. J., Duval, S. J., Tweedie, R. L., Abrams, K. R., & Jones, D. R. (2000). Empirical assessment of effect of publication bias on meta-analysis. *British Medical Journal*, 320, 1574-1577.
- Terrin, N., Schmid, C. H., & Lau, J. (2005). In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *Journal of Clinical Epidemiology*, 58, 894-901.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1-48.
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1, 80-83.

Footnote

¹ Duval and Tweedie (2000a, 2000b) describe two other estimators called R_0 and Q_0 . Their initial work found that the Q_0 estimator did not possess statistical properties as desirable as those of the L_0 and R_0 estimators. Research examining and employing the trim and fill procedure has largely focused on the L_0 estimator (Jennions & Møller, 2002; Terrin et al., 2003) and that choice has been maintained here. In particular, the initial demonstration of the technique by examining the publication bias rates in medical research by Duval and colleagues (Sutton et al., 2000) used the L_0 estimator. Duval (2005) notes that although both the L_0 and R_0 estimators tend to produce the same results, there are specific cases in which one produces estimates that are preferable to the other

² The Allen et al. (1996) and Cruz (1998) articles were not used from the Levine et al. (2009) sample because those articles did not contain meta-analyses of at least ten studies.

³ There are several other meta-analysis software programs that can produce trim and fill procedure outputs including comprehensive meta-analysis (Borenstein, Hedges, Higgins, & Rothstein, 2005) and MIX 2.0 (Bax, 2011).