2

PUBLICATION BIAS

Understanding the Myths Concerning Threats to the Advancement of Science

George C. Banks, Sven Kepes and Michael A. McDaniel

Concerns have been raised regarding the extent to which the scientific literature is free from bias (Banks & O'Boyle, 2013; Ioannidis & Doucouliagos, 2013; Kepes & McDaniel, 2013; Schmidt & Hunter, in press). In particular, researchers across scientific fields have been investigating the potential of publication bias as a threat to the advancement of science and evidence-based practice. Publication bias exists to the extent the literature that is publically available is not representative of completed studies on a particular relation of interest (Kepes, Banks, McDaniel, & Whetzel, 2012; Rothstein, Sutton, & Borenstein, 2005). In the organizational sciences, publication bias has most often been investigated by examining the potential for systematic differences between samples that appear in the published or readily accessible literature and those that do not (i.e., sample-level publication bias). In other fields, outcome-level publication bias (i.e., outcome-reporting bias), that is, the selective reporting of results, such as dropping unsupported hypotheses and related results from a manuscript (Kepes & McDaniel, 2013), has also been investigated (McGauran et al., 2010). Consistent with most disciplines, we use the term "publication bias" with reference to both sample-level and outcome-level publication bias.

Publication bias can result in distorted, typically inflated effect size estimates (Kepes et al., 2012; Schmidt & Oh, 2013) as well as theory proliferation (Leavitt, Mitchell, & Peterson, 2010). Both effects harm the advancement of science as well as evidence-based practice (Banks & McDaniel, 2011; Briner & Rousseau, 2011). There is evidence of sample-level publication bias across scientific areas, including medicine (Kicinski, 2013; Sterne et al., 2011; Sutton, 2009), psychology (Ferguson & Brannick, 2012; Kepes & McDaniel, 2013), education (Banks, Kepes, & Banks, 2012), economics (Ioannidis & Doucouliagos, 2013), and in the organizational sciences (Kepes et al., 2012; Kepes, Banks, & Oh, 2014; O'Boyle, Rutherford, &

Banks, in press; Vevea, Clements, & Hedges, 1993). Additionally, evidence of outcome-level publication bias has been documented in medicine (Chan & Altman, 2005; Dwan et al., 2008), sociology and political science (Gerber & Malhotra, 2008a, 2008b), psychology (Simonsohn, Nelson, & Simmons, 2014), and the organizational sciences (McDaniel, Rothstein, & Whetzel, 2006; O'Boyle, Banks, & Gonzalez-Mule, in press). A review of published scientific literature has indicated that the percentage of statistically significant results has increased by more than 22% from 1990 to 2007 (Fanelli, 2012). We suggest that it is likely that publication bias contributed to this increase in published statistically significant results and, in turn, that such results increase the likelihood of future publication bias.

Despite the growing body of publication bias research, several myths exist regarding publication bias. These myths relate to the operational definitions, causes, and approaches to detect and prevent this bias. This chapter reviews some of the myths, discusses the kernel of truth related to each myth, and then offers recommendations for future research.

Myth # 1: Publication Bias Is Concerned with the Availability of All Possible Effect Sizes in All Areas of a Scientific Field

In the typical case of publication bias, studies with small sample sizes and small magnitude effects are suppressed[1] from the published literature because the results are often not statistically significant. This suppression makes these studies difficult to locate when reviewing a literature (Hopewell, Clarke, & Mallett, 2005; Rothstein & Hopewell, 2009). Rothstein and colleagues (2005) defined publication bias as the extent to which "the research that appears in the published literature is systematically unrepresentative of the population of completed studies" (p. 1). To our knowledge, all publication bias studies prior to 2012 focused on relations between specific variables (e.g., an employment test and job performance, the Mozart effect on spatial ability). No one had attempted to declare that entire research literatures (e.g., organizational sciences) were free of or afflicted by publication bias.

In an attempt to investigate publication bias, Dalton, Aguinis, Dalton, Bosco and Pierce (2012) examined correlation matrices of published and unpublished studies and compared the statistical significance of the results in these two types of sources. Having found no difference, the authors stated, "we find that, contrary to the established belief, the file drawer problem is of little, if any, consequence for meta-analytically derived theoretical conclusions and applications in OBHRM [organizational behavior and human resource management], I-O psychology, and related fields" (p. 225). However, we observe a misconception in Dalton and colleagues' (2012) belief regarding what publication bias concerns and suggest that this limitation stems in part from the definition presented by Rothstein and colleagues (2005). Dalton and colleagues (2012) focused on the availability of all possible effect sizes in the management and I-O psychology literatures rather than effect sizes of focal relations. We observe that the Rothstein and colleagues (2005) definition did not explicitly specify the need to focus on focal relations. Consequently, we suggest that Dalton and colleagues (2012) misunderstood the publication bias literature and created confusion regarding how publication bias is operationally defined and assessed. Specifically, they created a myth that publication bias is concerned with the availability of all possible effect sizes in a scientific field rather than specific focal relations.

Kernel of Truth

The kernel of truth to this myth is that publication bias is often concerned with the availability of effect sizes based on their statistical significance. Yet publication bias is concerned with the availability of effect sizes on specific relations of interest (Banks & O'Boyle, 2013; Kepes & McDaniel, 2013). In fact, Dalton and colleagues (2012) acknowledged this limitation in their work and stated,

> we have not, however, established this phenomenon at the focal level. Our data do not provide an insight into whether such comparisons would maintain for studies—published and non-published—particularly focused on, for example, the 'Big Five' personality traits or employee withdrawal behaviors (e.g., absenteeism, transfers, and turnover).
>
> *(p. 244)*

Due to this limitation, they were unable, for example, to differentiate the variables from the correlation matrices that were dependent, independent, control, or moderator variables. This is important to consider because publication bias is more likely to emerge when some of the *hypothesized* relations are not found to be statistically significant (Kepes, McDaniel, Banks, Hurtz, & Donovan, 2011; Kepes, McDaniel, Brannick, & Banks, 2013). In sum, we assert that inferences cannot be made about the extent to which publication bias is or is not a problem based on the evidence provided by Dalton and colleagues (2012).

We note that meta-analytic reviews in the organizational sciences may not exclusively focus on relations that were explicitly of interest in the primary study. For example, Hunter and Schmidt (2004) explained that

> many meta-analyses focus on questions that were not central to the primary studies from which data are taken. For example, sex differences (in traits, abilities, attitudes, etc.) are rarely the central focus of a study; instead, they tend to be reported on an incidental basis, as supplementary analysis. Hence, these results tend not to be subject to publication bias because they are close to irrelevant to the central hypotheses.
>
> *(p. 497)*

As a result, there may be instances when publication bias is less likely to occur because the relations of interest for a meta-analysis were not the sole or major focus of the primary studies. Nonetheless, there is no existing evidence that indicates the frequency with which meta-analyses address questions that were not central to primary studies. Conversely, it is reasonable to assert that most meta-analyses are conducted on variables that were central to the hypotheses of primary research studies (Kepes et al., 2012).

Sorting Truth from Fiction

Publication bias involves the systematic suppression of effect sizes that are of interest to the research community. Consequently, the original definition of publication bias by Rothstein and colleagues (2005) may be refined to avoid confusion. Publication bias can be described as the extent to which research available to a reviewer is systematically unrepresentative of the population of completed studies on a specific relation of interest. Thus, investigations into the existence of publication bias studies should be focused on focal relations.

For example, Gerber and Malhotra (2008a, 2008b) examined the potential for an abundance of p-values just below the .05 threshold necessary to achieve statistical significance by the traditional standard. To accomplish this, these researchers coded p-values that were associated with hypothesized relations. Thus, they did not look at the statistical significance of variables that were not hypothesized to be correlated. The authors concluded that publication bias likely exists in the most methodologically rigorous journals within sociology and political science.

As another example, for 142 studies, O'Boyle, Banks, and Gonzalez-Mule (in press) investigated the chrysalis effect, which describes how dissertations undergo a metamorphosis from unpublished manuscripts to published journal articles. The authors focused specifically on hypothesized relations, and their results showed that the ratio of supported to unsupported hypotheses more than doubled (.82 to 1.00 versus 1.94 to 1.00) in the transition from a dissertation to journal article. This finding provided explicit and compelling evidence of outcome-level publication bias. Similarly, Bosco, Field, and Pierce (2012) found that mean correlations were noticeably larger when variable pairs were hypothesized to relate rather than not expected to relate. Combined, these studies demonstrate that publication bias clearly exists in the organizational science literatures, although it may not be present in all topics in the literature (Bosco et al., 2012; Gerber & Malhotra, 2008a, 2008b).

In meta-analytic reviews, sensitivity analyses should be used to estimate the extent of publication bias on relations of interest. A sensitivity analysis assesses the degree to which changes in analyses or included data influence results and conclusions (Greenhouse & Iyengar, 2009); publication bias analyses are best viewed as sensitivity analyses. To the degree that results and conclusions are not affected by publication bias, one can describe the results as robust. In the social sciences, publication bias findings have ranged from minimal or no bias (Banks et al., 2014; Chiaburu, Peng, Oh, Banks, & Lomeli, 2013; Harrison et al., in press; Kepes et al., 2012; Kepes et al., 2014) to more moderate and extreme cases of potential bias (e.g., Ferguson & Brannick, 2012; McDaniel, McKay, & Rothstein, 2006; McDaniel, Rothstein, & Whetzel, 2006; O'Boyle, Rutherford, & Banks, in press; Whetzel, 2006). For example, research that examined the relation between leader-member exchange (LMX) and team-member exchange (TMX) has shown little to no evidence of publication bias (Banks et al., 2014). Conversely, work in the field of entrepreneurship showed a strong likelihood of publication bias when examining the innovation-firm performance relationship (O'Boyle, Rutherford, & Banks, in press). Similarly, investigations into the possibility of publication bias in the natural sciences have always (to our knowledge) focused on specific relations of interest rather than the availability of all possible effect sizes (e.g., Blackwell, Thompson, & Refuerzo, 2009; Curfman, Morrissey, & Drazen, 2006; Song et al., 2010; Turner, Matthews, Linardatos, Tell, & Rosenthal, 2008).

In sum, the detection, evaluation, and prevention of publication bias should focus on relations of interest. The extent to which publication bias is a problem likely varies across research topics with large bias in some areas, moderate bias in others, and minimal or no bias in the remaining areas (Dickersin, 2005; Rothstein et al., 2005; Schmidt & Hunter, 2014). Hence the conclusion by Dalton and colleagues (2012) that "our results indicate that the methodological practice of estimating the extent to which results are not vulnerable to the file drawer problem may be eliminated" (p. 243) is clearly wrong. Sensitivity analyses to evaluate the presence and magnitude of publication bias are warranted (American Psychological Association, 2008, 2010; Borenstein, Hedges, Higgins, & Rothstein, 2009; Kepes et al., 2013; O'Boyle, Rutherford, & Banks, in press; Schmidt & Hunter, 2014). If the possibility of publication bias is investigated in a specific research topic, and no publication bias is found, we can have greater confidence in the robustness of the results. Such findings can only increase our confidence in the robustness of meta-analytic results and the trustworthiness of our cumulative scientific knowledge (Kepes et al., 2014; Kepes & McDaniel, 2013; McDaniel, Rothstein, & Whetzel, 2006).

Myth #2: The Editorial Review Process Is the Primary Cause of Publication Bias

There is also a myth related to the potential causes of publication bias. Specifically, there is a common misconception that editors and reviewers are the primary cause of publication bias. However, in the context of the medical literature, Dickersin (2005) wrote, "despite the consistent findings that only a small fraction of studies are not published because they are turned down by journals, investigators have persisted in naming bias at the editorial level as the main reason why negative or null results are not published" (p. 21; see also Chalmers & Dickersin, 2013). Hence, authors may be the primary cause of publication bias.

Kernel of Truth

There is a kernel of truth to this myth, as it is likely that editors and reviewers may have predispositions to reject manuscripts that contain mixed or statistically nonsignificant results. Evidence suggests that reviewers in the social sciences may be biased by positive results in that they are more likely to recommend that a manuscript with statistically significant findings be published compared to the same manuscript without statistically significant findings (e.g., Epstein, 1990; Mahoney, 1977). Anecdotal evidence also suggests that rejection letters often state that papers are being rejected because the hypotheses were not supported. Recently, Emerson and colleagues (2010) found similar results and illustrated that reviewers were more likely to be critical of research methods and to find purposefully planted errors within a manuscript when the results were negative. Thus, the body of evidence indicates that editors and reviewers do play a role in the existence of publication bias.

Additionally, investigations indicate the existence of a type of bias referred to as *outlet bias*. This bias can be described "as occurring when the place of publication is associated with the direction or strength of the study findings" (Song et al., 2010, p. 3). It appears that studies submitted to higher-quality journals may be more likely to be accepted if they contain a high percentage of supported hypotheses (i.e., statistically significant results). Findings from the natural sciences, such as studies in ecology and medicine, suggest that primary studies with predominantly statistically significant results are more likely to be published in higher-impact journals (Easterbrook, Gopalan, Berlin, & Matthews, 1991; Etter & Stapleton, 2009; Murtaugh, 2002).

However, there is also evidence in the medical literature that editors and reviewers may not be responsible for much of the publication bias. One study considered 745 manuscripts submitted to the *Journal of the American Medical Association* (Olson et al., 2002). The investigators concluded that there was no meaningful difference in the likelihood of publication between those studies with positive findings compared to those with negative results. Although there is some evidence that both editors and reviewers contribute to the existence of publication bias, it appears that authors are the primary cause of publication bias (Chalmers & Dickersin, 2013; Dickersin, 2005).

Sorting Truth from Fiction

Research indicates that authors are more likely to submit their studies to a journal if the findings are statistically significant (Kepes & McDaniel, 2013; Schmidt & Hunter, 2014). Clearly, authors have a greater opportunity to engage in practices that result in publication bias prior to the peer-review process (Chalmers & Dickersin, 2013). Banks and McDaniel (2011) detailed nine reasons that authors or organizations may not want to submit a study to the peer-review process (see also Kepes et al., 2014). For instance, the findings of one's study might not be statistically significant or the results may be contrary to theory or past findings. Authors may not submit such a study and instead submit other studies that they believe have a greater chance of acceptance in a journal (i.e., studies with statistically significant results).

When authors are focused on submitting studies with results they believe have the best chance of being published (i.e., studies with significant results; Hartshorne & Schachner, 2012; Sterling & Rosenbaum, 1995), the authors cause publication bias by not submitting manuscripts with statistically nonsignificant results. Additionally, because organizational researchers typically test multiple hypotheses using multiple variables, authors may only report those with statistically significant findings (Kepes & McDaniel, 2013; O'Boyle, Banks, & Gonzalez-Mule, in press). For example, McDaniel, Rothstein, and Whetzel (2006) found that an employment test vendor, with a test product that yields multiple scale scores, reported the validity for some scales for a given sample but not for other scales for the same sample. This reporting practice is consistent with an inference of outcome-level publication bias designed to suppress small-magnitude results.

Evidence from psychological research also indicated that authors are likely to be the primary cause of publication bias (e.g., Greenwald, 1975). Cooper, DeNeve, and Charlton (1997) examined author behavior concerning 117 completed studies. Of these studies, approximately 62% had statistically significant results and 50% of these were submitted for peer review to conferences compared to just 7% of the studies with statistically nonsignificant findings. In terms of submissions for publication, those with statistically significant results were submitted 74% of the time compared to just 4% for studies with statistically nonsignificant findings. Consequently, it is clear that authors often engage in practices that result in publication bias.

In summary, evidence suggests that editors and reviewers are one cause of publication bias (Emerson et al., 2010; Epstein, 1990; Mahoney, 1977). However, authors are likely to be the primary cause of this bias (Chalmers & Dickersin, 2013; Cooper et al., 1997; Dickersin, 2005; Greenwald, 1975) because they have control over their data and decide whether a manuscript based on that data is submitted to a journal (Banks & McDaniel, 2011). If reviewers elect to reject a study because the results are not statistically significant, authors can resubmit their study to another journal or attempt to find other means to disseminate their results. Authors likely make the accurate assumption (Kepes & McDaniel, 2013) that editors and reviewers have a preference for positive results because such results are newsworthy (Dickersin, 2005). Thus, papers with mostly statistically nonsignificant results may never be disseminated, yielding publication bias and an overestimation of effect magnitude in our published literature.

Myth #3: The Failsafe N and Subgroup Comparisons Are the Best Publication Bias Detection Techniques

Until now, we have explored myths stemming from the definition of publication bias and the potential causes of this bias. We now consider myths related to how publication bias may be detected and prevented. There are numerous tests that can be used to detect the potential presence and influence of publication bias. However, evidence suggests that from 2005 to 2010, only 31% of meta-analytic reviews in the organizational sciences tested for the possibility of publication bias (Banks, Kepes, & McDaniel, 2012; Kepes et al., 2012). Furthermore, when such tests were performed, the failsafe N and the subgroup comparison by data source (e.g., published vs. unpublished) were the two most popular techniques. However, neither of these methods allows for an adequate assessment of publication bias (Becker, 1994, 2005; Evans, 1996). In brief, the majority of meta-analytic reviews in the organizational sciences do not consider the threat of publication bias and, when considered, authors use methods that are arguably the least effective at evaluating the presence of this bias and do not assess the magnitude of the bias.

Given the widespread use of the failsafe N and subgroup comparisons by data source, there appears to be a belief that these methods are adequate techniques to assess the potential threat of publication bias. There are some kernels of truth behind this belief. After a brief review of these truths, we will present a summary of the current state of the literature that includes a discussion of recommended techniques for evaluating the presence of publication bias in meta-analytic reviews.

Kernel of Truth

The failsafe N was first introduced by Rosenthal (1979) as a means to examine the possibility of publication bias. The failsafe N procedure purports to estimate the number of statistically nonsignificant effect sizes that would be needed to make a statistically significant meta-analytic mean effect size statistically nonsignificant. Hence, the kernel of truth is that the failsafe N was the first offered technique to detect the presence of publication bias. The method was offered to address the concern that the conclusion of a meta-analytic review could change if a large number of missing, statistically nonsignificant results were obtained and added to the review. Thus, Rosenthal proposed the following question: If a statistically significant meta-analytic result was obtained, how many more effect sizes would be needed to reduce it to a point of statistical nonsignificance (i.e., nullify the result)? If the number of additional effect sizes was small, one might have cause for concern. If the number was large, one could be more confident that the meta-analytic mean effect size magnitude was not meaningfully influenced by publication bias.

McDaniel and colleagues (2006) provided an example of the limited utility of the failsafe N using a scenario in which an employer must choose between two different employment selection tests (Tests A and B), each measuring the same construct. For Test A, the mean validity is .25 compared to a mean validity of .20 for Test B. McDaniel and colleagues (2006) explained that, all else being equal (cost, ease of administration, etc.), the employer might select to use Test A if one assumes no publication bias. However, as noted by McDaniel and colleagues, "Knowing that it takes 80 file drawer studies to nullify the validity of Test A and

100 file drawer studies to nullify the validity of Test B does not help to determine the validity of the tests in the absence of publication bias" (2006, p. 930). As such, the failsafe N provides no useful information concerning the validity of the two tests and does not inform the decision concerning which test to use.

Several reviews have described the substantial limitations of the failsafe N (Becker, 1994, 2005; Evans, 1996). Here, we draw largely on Becker (2005). The first limitation is that the failsafe N is based on the improbable assumption that all missing effect sizes are zero in magnitude. Second, the failsafe N focuses on the statistical significance of meta-analytic estimates and ignores the magnitude of the effect size, which is of greater importance. Third, different approaches yield widely varying estimates of the failsafe N. Fourth, no statistical criteria are available to assist in interpretation. Fifth, the failsafe N does not incorporate sample size information. To address limitations of the failsafe N, modifications have been offered (Orwin, 1983), but these modifications do not improve the effectiveness of the technique substantially (Becker, 2005; Higgins & Green, 2011). In sum, Becker (2005, p. 111) suggested that "the failsafe N should be abandoned in favor of other more informative analyses." Despite all the evidence against the use of failsafe N, editors and reviewers in organizational science journals often recommend its use.

Similarly, the subgroup comparison has been used as a means to evaluate the presence of publication bias. For instance, Schmidt and colleagues (1985) compared published and unpublished samples to determine if published samples reported more statistically significant results than those samples that were unpublished. Similarly, Lipsey and Wilson (1993) and McKay and McDaniel (2006) reported mean effect size estimates by data source. Thus, the kernel of truth concerning the use of the subgroup comparison is that one can sometimes find mean effect size differences between published and unpublished samples. The limitation of a subgroup comparison is that it only provides an estimate of the difference between *identified published* and *identified unpublished* samples. Thus, the analysis is based on the assumption that all samples with relevant effect sizes, published and unpublished, have been identified and included in the meta-analysis. This is improbable in the social sciences (Hopewell et al., 2005; Kepes et al., 2012) and particularly improbable for unpublished samples (Ferguson & Brannick, 2012). Finally, publication bias can be in the same direction or in opposite directions in effect size distributions subset by data source (Kepes et al., 2014).

Sorting Truth from Fiction

The truth is that the failsafe N should never be used for publication bias analyses, and the comparison by data source is a suboptimal test of publication bias. We do not discourage the use of data source comparisons (e.g., published vs. unpublished), but the analyses should be supplemented with methods that permit clearer inferences about the extent and magnitude of publication bias (e.g., Banks et al., 2014; Kepes et al., 2012). More advanced methods, used in combination to triangulate the results, are recommended (Kepes et al., 2012; Kepes et al., 2014). These methods include contour-enhanced funnel plots (Palmer, Peters, Sutton, & Moreno, 2008; Peters, Sutton, Jones, Abrams, & Rushton, 2008; Sterne et al., 2011), trim and fill (Duval, 2005; Duval & Tweedie, 2000a, b), cumulative meta-analysis (Borenstein et al., 2009; McDaniel, 2009), and selection models (Hedges & Vevea, 2005; Vevea & Woods, 2005). We describe each in the following sections.

Funnel Plots

A funnel plot is used to illustrate graphically the magnitude of an effect size plotted along an X axis relative to precision (inverse of a sample's standard error) presented along a Y axis (see Figure 2.1). In homogenous distributions of effect sizes (i.e., distributions in which effect size variation is solely due to random sampling error), effect sizes from large samples have less random sampling error (i.e., greater precision) on average and tend to cluster at the top of the funnel plot around the estimated mean. Conversely, smaller samples with greater random sampling error (i.e., less precision) tend to vary more widely. In the absence of heterogeneous variance (e.g., variance due to moderators), the distribution of samples in the funnel plot will be symmetrical (see Figure 2.1a; Sterne, Gavaghan, & Egger, 2005). In the event that publication bias has suppressed small-magnitude effects, small samples with results that are statistically nonsignificant will be absent, resulting in an asymmetric distribution (see Figure 2.1b).

Funnel plot asymmetry due to heterogeneity (i.e., variance not due to random sampling error) is possible and may distort conclusions concerning publication bias. Any form of heterogeneity can cause problems when drawing inferences concerning publication bias analyses, but a heterogeneity cause (e.g., moderator) that covaries with sample size is particularly problematic. For example, small-sample studies may use more reliable measures (e.g., biological markers of strain associated with work stress) yielding larger-magnitude effects, but large-sample studies may rely on self-report measures, yielding smaller-magnitude effects. In this scenario, effect size will likely be correlated with sample size, resulting in funnel plot asymmetry that may be mistaken for publication bias. Thus, it is recommended that publication bias tests be used within more homogeneous subgroups or that techniques such as meta-regression be used (Kepes et al., 2012).

Although the suppression of small-magnitude effects is likely to be the most common scenario in data affected by publication bias, asymmetry may be due to suppression of large-magnitude effects. In these circumstances, effect size suppression may occur when large-magnitude effects are socially uncomfortable to report, such as with age, race, or sex differences (McDaniel, McKay, & Rothstein, 2006; Tate & McDaniel, 2008). To help differentiate asymmetry due to heterogeneity from asymmetry due to publication bias, one can use contour-enhanced funnel plots. These plots incorporate contour lines that correspond to commonly used values of statistical significance (i.e., $p < .05$ and $p < .10$), which aids in

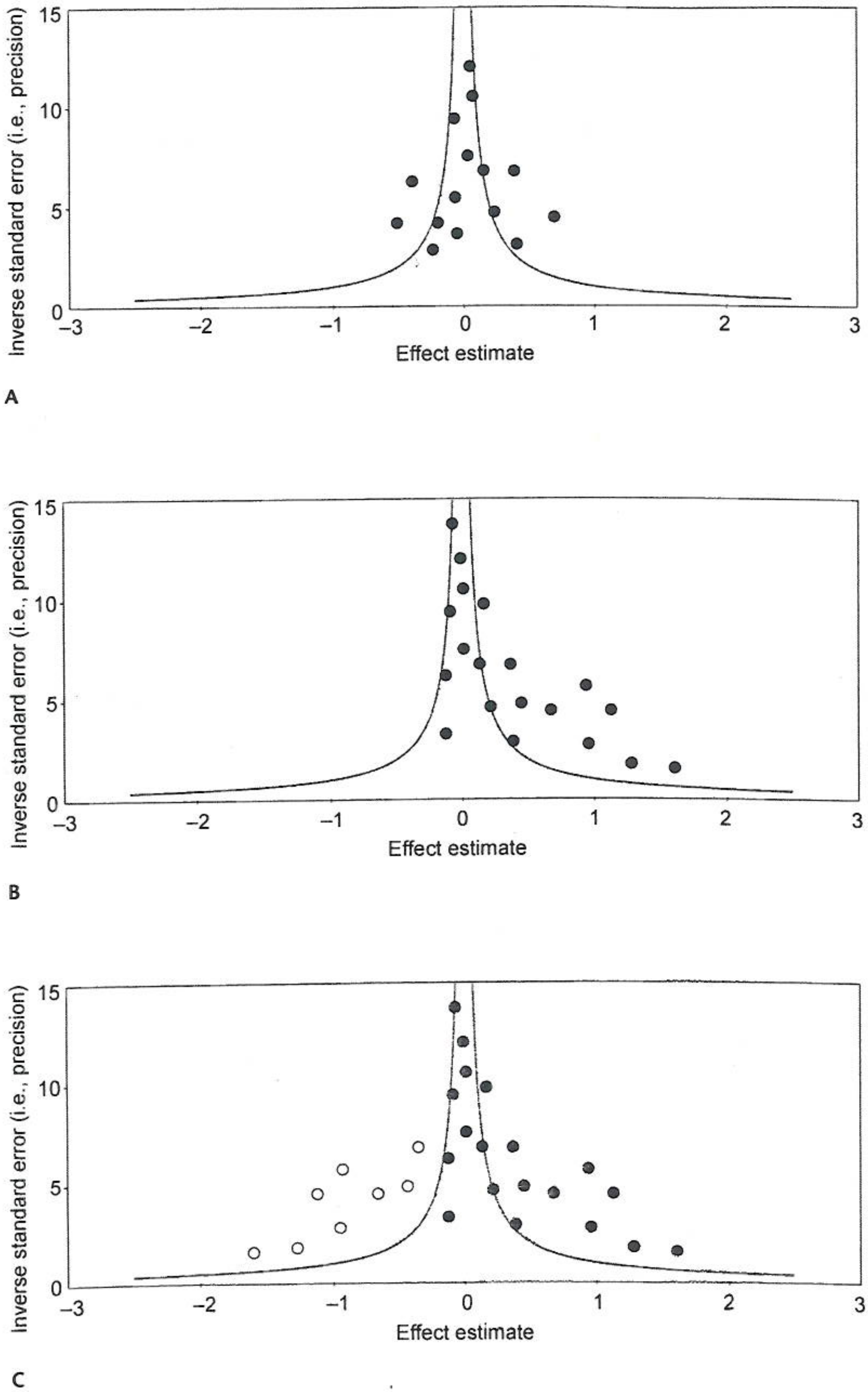

FIGURE 2.1 Exemplar funnel plots

distinguishing publication bias from other causes of funnel plot asymmetry (Kepes et al., 2012; Peters et al., 2008; Sterne et al., 2011).

Trim and Fill

The trim and fill technique evaluates the symmetry of funnel plot distributions and seeks to estimate the magnitude of the mean effect if the assumed suppressed studies were present. When the funnel plot distribution is asymmetric (e.g., small-sample, small-magnitude effect sizes are missing from a distribution), the trim and fill procedure "trims" effect sizes from the nonskewed side in the funnel plot in an iterative approach until a symmetrical distribution is achieved. A new mean based on this trimmed distribution is calculated. Next, trimmed effect sizes are returned to the distribution and a set of imputed effect sizes is added to the distribution to achieve symmetry (see Figure 2.1c) around the new mean. At the conclusion of this process, one has the mean of the original effect size distribution and the mean of a distribution containing both the original effect sizes and the imputed effected sizes (sometimes called the trim and fill adjusted distribution). To the extent that the two means are different, one can infer the degree or magnitude of the publication bias.

Kepes and colleagues (2012) developed decision rules that can be adopted to judge the importance of the magnitude of the difference in means (McDaniel, Rothstein, & Whetzel, 2006; Rothstein et al., 2005). For example, small or no differences between the original meta-analytic estimates and the adjusted estimates may be interpreted as minimal to no evidence for publication bias. Specifically, one might declare that less than a .05 absolute change and less than a 20% relative change in the mean estimates suggest that publication bias is at most minimal. One might interpret moderate publication bias if there is at least a .05 absolute change and more than a 20% relative change but less than 40% in the meta-analytic estimate. Finally, if there is a large-magnitude difference between the original meta-analytic estimate and the adjusted estimate, one might conclude that an extreme case of publication bias exists. One might interpret that there is a large degree of publication bias if there is at least a .05 absolute change and at least a 40% relative change in the meta-analytic estimate. These or other decision rules could be used for specific research efforts (see Kepes et al., 2012; Kepes et al., 2014).

Some cautions are necessary concerning trim and fill analyses. Given the effect size imputation, it is unwise to interpret the mean of the trim and fill adjusted distribution as the "true" estimate of the mean effect. Rather, one compares the difference between the two means as a sensitivity test to judge the likelihood of bias stemming from effect size suppression. Trim and fill also assumes that the effect sizes are homogeneous (e.g., no moderators are present), and the method is not robust to violations of this assumption (Terrin, Schmid, Lau, & Olkin, 2003). Trim and fill would best be used on subdistributions in which moderators are largely controlled. One could also consider a meta-regression procedure (Weinhandl & Duval, 2012) that controls for variance due to moderators and then applies trim and fill to the residuals.

Also, trim and fill can be combined with the contour-enhanced funnel plot. If the imputed effect sizes are not statistically significant, one can infer publication bias due to suppression of statistically nonsignificant effect sizes. If the imputed studies are statistically significant, one should look for a small sample effect (Kepes et al., 2012; Kepes et al., 2014; Peters et al., 2008).

Selection Models

Selection models (Hedges & Vevea, 2005; Vevea & Woods, 2005) are another technique to detect publication bias and have a different set of assumptions than methods that rely on funnel plot symmetry. Selection models allow one to examine how meta-analytic results may be affected by selection processes that are influenced by study characteristics, typically statistical significance. When a meta-analysis does not consider the possibility that publication bias is present, one is making the assumption that one has 100% of all extant effect sizes in the meta-analysis and, thus, that no bias is present. Selection models, as typically applied, assume that the probability of a sample being observed (i.e., included in the meta-analysis) depends to some extent on the statistical significance of the effect size. Selection models then reweight effect sizes based on their statistical significance.

In the organizational sciences, a priori selection models are typically used (Hedges & Vevea, 2005; Kepes et al., 2012; Vevea & Woods, 2005). As an example, when operating under the assumption of moderate publication bias, an effect size with a *p*-value in the range of .000 to .005 can be assigned a 100% probability of being included in a meta-analysis (a weight of 1.0). Conversely, an effect size with a *p*-value that falls within the range of .500 to .650 might only have a 60% probability of being observed and hence would be assigned a weight of .60 (for a complete list of proposed weights under moderate and severe assumptions of publication bias, see Table 5 in Vevea & Woods, 2005). After assigning weights to effect sizes based on their statistical significance, an adjusted meta-analytic effect size estimate is calculated. As with the trim and fill analysis, the difference between the original adjusted mean estimates allows inferences regarding the potential effects of publication bias on the original mean estimate. Thus, the same decision rules can be applied to judge the degree of publication bias (see Kepes et al., 2012). Vevea and Wood (2005) reported that selection models are relatively robust to effect size heterogeneity.

Cumulative Meta-analysis

Cumulative meta-analysis is a technique that requires one to sort a set of effect sizes based on a characteristic of interest. A set of meta-analyses is then calculated in an iterative process whereby effect sizes are added one at a time to the analysis, each time calculating a new mean estimate. When used as a publication bias analysis, one sorts effect sizes by precision as illustrated in Figure 2.2. Thus, the most precise

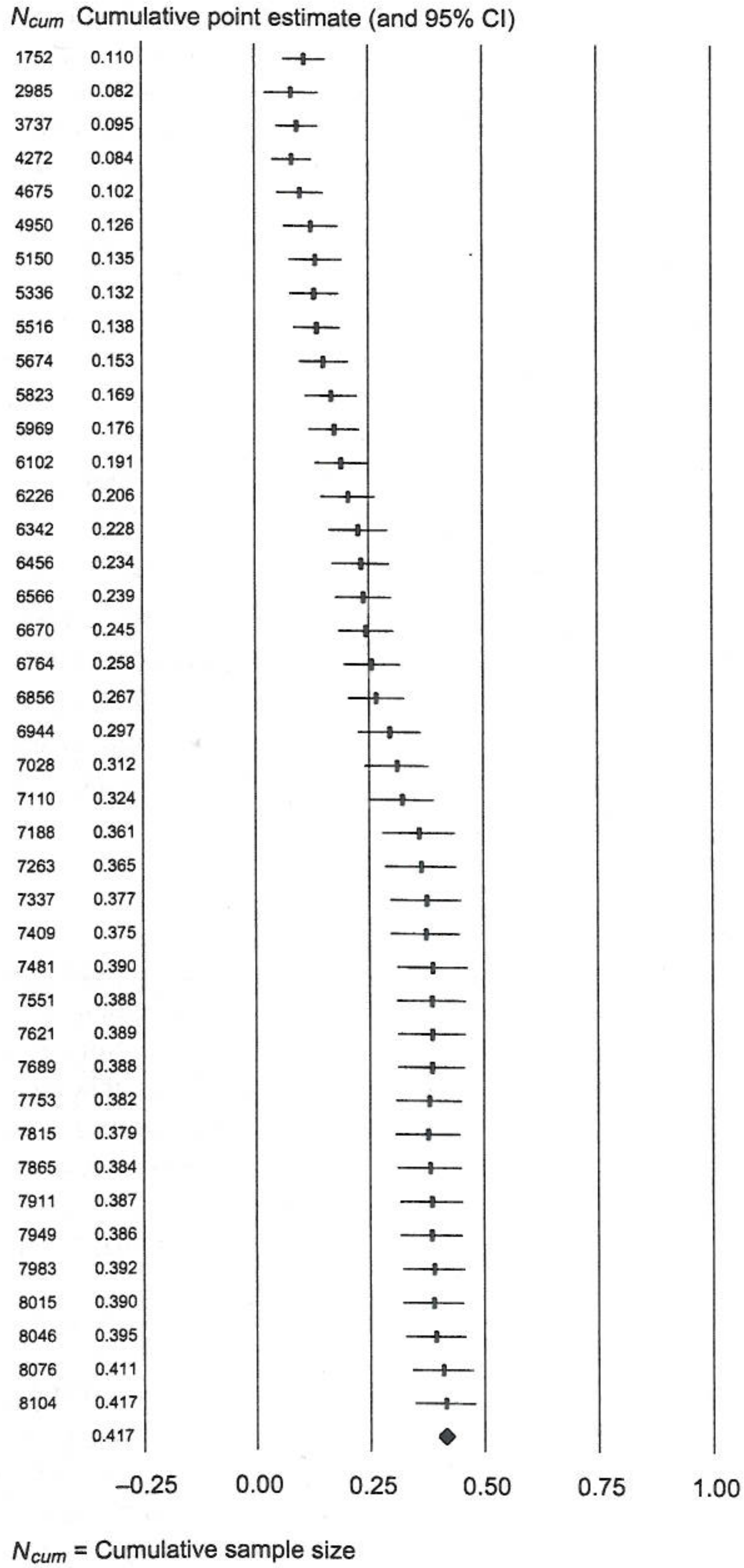

FIGURE 2.2 Exemplar cumulative meta-analysis

sample is added first, followed by the second most precise sample and so forth. The mean effect size estimates calculated at each step can be illustrated in a forest plot. Evidence of "drift" in a forest plot is consistent with an inference of publication bias. In the typical case of publication bias in which small-sample, small-magnitude effect sizes have been suppressed, the plot of cumulative mean estimates drifts to the right (the mean effect size gets larger as less-precise effects are iteratively added to the distribution). Decision rules are needed to determine the severity of the drift (Kepes et al., 2012). For example, one might compare the five most precise samples to the final cumulative mean estimate or the cumulative mean estimate of the most- and least-precise samples (e.g., the 25% most- and least-precise samples; Kepes et al., 2012) and examine the magnitude of the difference. Alternatively, one can compare the cumulative means at fixed intervals (e.g., every 10th mean) or at fixed intervals of cumulative sample size (e.g., compare the means at cumulative samples based on 5,000 observations versus 10,000 observations).

Other Methods

Other publication bias methods are available, including Egger's test of the intercept (Egger, Smith, Schneider, & Minder, 1997) and Begg and Mazumdar's (1994) rank correlation test. Yet, because of their low statistical power and related statistical concerns, Kepes and colleagues (2012) did not recommended their use for every examination of publication bias.

Triangulation

As with all statistical techniques, the methods to detect and adjust for publication bias are only as good as their underlying assumptions. If assumptions are inaccurate, any method may reach erroneous results (Ioannidis, 2008). Nonetheless, if one does not use sensitivity analyses to test for the potential of publication bias, one runs the risk of making the false assumption that publication bias is not an issue and this assumption is not tested in any form (Vevea & Woods, 2005). Consequently, we recommend that researchers test for publication bias in all meta-analyses using multiple publication bias detection techniques that rely on different assumptions.

The use of multiple publication bias methods is a form of triangulation (Kepes et al., 2012; Kepes et al., 2014). Triangulation can be defined as the use of "multiple reference points to locate an object's exact position" (Jick, 1979, p. 602). In management research, triangulation characterizes the use of multiple study designs, settings, samples, and methods to investigate a particular phenomenon (Sackett & Larson, 1990). The use of multiple publication bias methods triangulates conclusions concerning publication bias in that one obtains multiple publication bias results using methods that do not necessarily share the same assumptions (Ferguson & Brannick, 2012; Kepes et al., 2012; Kepes et al., 2014). For example, some methods use asymmetry to inform inferences of publication bias, while other methods use the magnitude of mean differences or drift in cumulative meta-analysis to inform inferences.

Future research should attempt to better understand the influence of artifactual variance (e.g., measurement error, range restriction) as well as outliers on the robustness of the various publication bias techniques. Relatedly, more research is needed to address the effect of heterogeneity in publication bias analyses. In addition, more work is needed to examine the effects of outcome-level publication bias on meta-analytic estimates (Biemann, 2013; Hahn, Williamson, Hutton, Garner, & Flynn, 2000; Kirkham, Riley, & Williams, 2011; Williamson & Gamble, 2007). However, the best approach to mitigate publication bias is to engage in practices to prevent publication bias from occurring in the first place. We conclude this chapter by discussing a myth related to the prevention of publication bias.

Myth #4: Publication Bias May Not Be of Concern in the Social Sciences because It Is Prevented by Reporting Correlation Matrices, Testing Multiple Hypotheses, and Conducting Systematic Searches in Meta-analytic Reviews

Despite the fact that empirical evidence for the presence of publication bias has been found in virtually all scientific fields where it has been investigated, it has been suggested that publication bias may not be as much of a concern in the organizational sciences (Dalton et al., 2012; Hunter & Schmidt, 2004). Based on the publication bias research conducted to date, Schmidt has changed his position on publication bias and now considers it a topic that should be addressed (Schmidt & Hunter, 2014). The assertion that publication bias is not a concern in the organizational sciences is in part drawn from the observation that studies published in organizational science journals typically report correlation matrices and test multiple hypotheses. The argument is also based on assertions that the results are less likely to depend on the statistical significance of an individual outcome and that unpublished studies can be identified through systematic searches. Thus, it is assumed that statistically nonsignificant effect sizes are still available in the publically available literature and, as a result, a myth has emerged that publication bias may not be of concern in the social sciences.

Kernel of Truth

The kernel of truth is that some practices can reduce the potential for publication bias. Correlation matrices allow researchers to report multiple relations, and, in particular, relations that may not have been the main focus of a primary study. However, many, perhaps most, meta-analyses draw data from studies where the research question addressed in the meta-analysis is a central part of "interesting" hypotheses in primary studies (Kepes et al., 2012), such as job satisfaction, personality traits, transformational leadership, leader-member exchange (LMX), and predictors of individual- and firm-level performance. Thus, assertions that meta-analyses concern relations between variables tangential to the main focus of the primary studies are often likely incorrect (Kepes et al., 2012).

Additionally, in the organizational sciences, researchers typically test multiple hypotheses, unlike, for instance, in medical research. This practice makes it less likely that authors are dependent upon the success of any one hypothesis. Hence, researchers may not have their paper rejected because of a lack of support for any one particular hypothesis. When testing multiple hypotheses, at least some hypotheses may be supported by chance. As a result, it may be uncommon that a researcher does not have any statistically significant findings to report. Testing of multiple hypotheses may make it easier for some null results to be published, but it may not completely eliminate publication bias. For example, O'Boyle and colleagues (in press) found that recent doctoral graduates showed a strong preference for eliminating unsupported hypotheses from their dissertation studies and a strong preference for adding supported hypotheses post hoc before submitting their dissertation study for publication. These researchers also appeared to engage in practices such as adding and deleting data as well as adding and removing variables in order to turn unsupported hypotheses into supported ones. Additionally, researchers sometimes collect data using multiple operationalizations of the same variables and may only report the variables and relations that were statistically significant. These examples illustrate outcome-level publication bias. When unsupported hypotheses are dropped and others are added post hoc with a preference for statistically significant results, biased correlation matrices get published (Biemann, 2013). Consequently, neither the use of correlation matrices nor the practice of testing multiple hypotheses is likely to prevent publication bias, although they may reduce it to some degree.

Additionally, it is asserted that a thorough systematic search can be used to identify unpublished studies that should be included in a meta-analytic review. By identifying and including effect sizes from unpublished studies, it has been suggested that meta-analytic reviews can overcome the threat of publication bias (Hopewell et al., 2005; Hunter & Schmidt, 2004; Rothstein & Hopewell, 2009). Thus, identifying unpublished samples (i.e., their effect sizes) and including them in a meta-analytic review can be considered a best practice that should be encouraged (Rothstein, 2012). However, it is difficult to conclude that even the most thorough systematic searches will identify all unpublished samples or a representative sample of them. For example, Ferguson and Brannick (2012) observed that unpublished samples in meta-analyses are often from the authors that conducted the meta-analysis. Although it is appropriate to include such unpublished samples, it is not credible to suggest that a meta-analytic study contains all unpublished samples or that the identified and included samples are representative of all unpublished samples. Also, even when relevant unpublished samples can be identified, the authors of such samples may be unable or unwilling to share the studies (e.g., Banks, Batchelor, & McDaniel, 2010; Banks et al., 2014; McDaniel & Kepes, in press).

Sorting out Truth from Fiction

A systematic search is unlikely to identify *all* samples with relevant effect sizes (Hopewell et al., 2005), nor do the reporting of correlation matrices and testing of multiple hypotheses eliminate the suppression of entire samples or even individual outcomes (Sutton & Pigott, 2005). Correlation matrices, testing multiple hypotheses, and systematic searches of the literature are means to reduce the potential presence of publication bias. However, such steps cannot completely prevent this bias from being a concern in the organizational sciences. In the last section of this chapter, we briefly describe steps that can be implemented to prevent such publication bias from occurring.

Recommendations for Preventing Publication Bias

Honor Codes

Honor codes are a potential means to reduce outcome-level publication bias (i.e., outcome-reporting bias). If journals were to implement an honor code, they would ask authors submitting a paper for review to sign a statement acknowledging that they did not engage in any questionable research practices (QRPs; e.g., O'Boyle, Banks, & Gonzalez-Mule, in press). In the event that authors did engage in QRPs, they would have the opportunity to acknowledge the QRP and disclose the practice and the logic behind it. QRPs include adding and dropping hypotheses from a study with a preference for those with statistically significant results. QRPs also include adding data (to increase statistical power) and dropping data (e.g., removing outliers), altering data, and adding and dropping variables, as well as hypothesizing after the results are known (Kerr, 1998). QRPs could range from more benign in nature, such as not reporting all dependent variables collected in a study, to clear ethical violations, such as falsifying data. The use of the term "QRP" is meant to suggest that although such practices may be questionable, they may not be unethical because their use may sometimes be appropriate but at other times inappropriate, depending on the specific practice and context (e.g., the deletion of outliers). However, it is clear that QRPs can result in outcome-level publication bias (O'Boyle et al., in press).

Researchers could also be asked to disclose other variables collected or investigated, as it is often common in observational studies in the social sciences to collect as much data as possible. The Journal Article Reporting Standards (JARS) of the American Psychological Association (2008) recommend such a disclosure. Honor codes can serve to prevent researchers from engaging in QRPs out of ignorance and would give journals the ability to more easily retract an article or print an erratum should evidence emerge that authors engaged in QRPs that led to publication bias. Unfortunately, such cases do occur in the field of management, as illustrated by the recent retractions issued by some of our top journals, including the *Academy of Management Journal*, *Journal of Management Studies*, *Journal of Business Venturing*, *Organization Science*, and *Strategic Management Journal* for the work completed by Ulrich Lichtenthaler. The practice of implementing such codes of conduct has proven to be effective for the reduction of some questionable practices (Ariely, 2012; Mazar, Amir, & Ariely, 2008).

Supplemental Information on the Internet

Providing supplemental information would allow for the dissemination of additional materials about one's study design, the study population, and any analyses and results not included in the submitted or published article. It is possible that reviewers and editors recommend the removal of such information because they do not find this information to be necessary or informative or because of space constraints. Online supplemental information would provide authors and journals with the opportunity to make this information widely available at little or no cost. As noted by Kepes and McDaniel (2013), if the Internet has room for millions of cat videos, it has room for such supplemental information. This should serve to reduce concerns due to publication bias (Wertil, Schob, Brunner, & Steurer, 2013).

Data Sharing

Journals should consider requesting that authors submit their raw data along with their manuscript (Kepes & McDaniel, 2013). This practice would increase editors' and reviewers' confidence in the analysis, as they would be able to analyze the data themselves should they have any questions. Additionally, journals could make the data available on their websites at the time of publication or after a grace period (e.g., 3 years). Thus, other researchers would have the opportunity to replicate the results and include the data in a meta-analytic review more easily (Kepes & McDaniel, 2013).

Two-stage Review Process

Kepes and colleagues (Kepes, Bennett, & McDaniel, 2014; Kepes & McDaniel, 2013) also suggested the implementation of a two-stage review process to minimize the use of QRPs. Editors and reviewers are biased toward the publication of articles with statistically significant results (e.g., Epstein, 1990; Mahoney, 1977) and are less critical of a study's methods when most results are positive (Emerson et al., 2010). Thus, blinding them to a study's Results and Discussion sections during an initial stage of the review process could minimize the introduction of these and related biases in the editorial review process. In the first stage, reviewers could be presented with the

Introduction and Methods sections, including a detailed description of the analysis approach, and the editor would receive comments and ratings free from any bias that might be created due to the support or lack of support from the results. Based on the comments from the reviewers, the author(s) would revise the paper and resubmit it. This submission could include the Results and Discussion sections. Then the reviewers would only have to check if the authors actually followed their previously submitted and reviewed methods and analysis plan and provided an appropriate discussion of results.

Incentives for B- and C-tier Journals

Universities should consider providing incentives for publications in B- and C-tier journals. It is not uncommon for universities, particularly for top research universities, to provide incentives only for A-tier journal publications. This incentive system may discourage authors from working on manuscripts that have been rejected from the best journals and may perpetuate publication bias (Kepes & McDaniel, 2013). Hence, publication bias could be mitigated if researchers were still provided with at least some sort of incentive for disseminating their studies in outlets other than A-tier journals.

Journal Submission and Communications

Another suggestion is the release of original journal submissions as well as communications among action editors, reviewers, and authors (Kepes & McDaniel, 2013; O'Boyle, Banks, & Gonzalez-Mule, in press). Editors, reviewers, and authors could be made aware in advance that their communications will be publically available. This might discourage researchers from engaging in QRPs once the review process has begun, and it might discourage editors and reviewers from encouraging the engagement in QRPs (e.g., presenting post hoc hypotheses as a priori).

Replications and Prospective Meta-analysis

Another recommendation for reducing issues related to publication bias is to encourage more replication studies. Journals should consider dedicating space solely for exact and conceptual replication studies. Only limited journal space would be required to publish replication studies because a literature review and theoretical framework would not need to be described (Kepes & McDaniel, 2013; O'Boyle, Banks, & Gonzalez-Mule, in press).

Additionally, prospective meta-analyses can be implemented as means to encourage simultaneous replication studies (Berlin & Ghersi, 2005). Prospective meta-analyses involve a group of researchers who collaborate to collect multiple samples in an investigation of the same research questions. Such an approach would allow for the standardization of research designs (e.g., measures) and the a priori inclusion of moderating variables that should be considered across the different research teams and samples. It would thus allow researchers to simultaneously replicate findings. Additionally, a prospective meta-analysis allows for the triangulation of results because of the use of multiple measures and design approaches. Finally, because prospective meta-analyses are planned prior to data collection, results tend to be reported regardless of whether they are statistically significant. Hence, this approach should serve to reduce publication bias.

Study and Protocol Registration

Numerous calls have been made for the registration of studies prior to completion as well as the creation of research registries for studies that have been completed but have not been published (Banks, Kepes, & Banks, 2012; Banks, Kepes, & McDaniel, 2012; Banks & McDaniel, 2011; Bennett & Miao, 2013; Ferguson & Brannick, 2012; Kepes et al., 2012; Kepes et al., 2014; Kepes & McDaniel, 2013). Additionally, incomplete research registries have plagued the medical field, where they have already begun to be implemented (Chan, 2008). However, following the mandatory registration requirement for medical trials (De Angelis et al., 2004), there was a substantial increase in the number of registrations. Furthermore, the registration data were more complete. Thus, the registration requirement resulted in less data suppression and more publically available data (Zarin, Tse, & Ide, 2005; Zarin, Tse, Williams, Califf, & Ide, 2011), which should yield more accurate mean effect size estimates in meta-analytic reviews. Therefore, anything short of mandatory participation in registries may still lead to a biased meta-analytic samples in which there are systematic differences between those researchers that willingly participate in the registries and those that do not (Strech, 2011).

Conclusion

Publication bias can present a serious threat to the advancement of science. Publication bias has been documented in several investigations in the organizational sciences. As conscientious researchers, we should evaluate the extent to which this bias exists within individual literature areas and take proactive steps to prevent it from occurring. The recommendations discussed throughout this chapter are summarized in Table 2.1.

In this chapter, we reviewed myths related to publication bias. Specifically, we clarified that publication bias is concerned with systematic differences between the literature that is available to a reviewer and the population of completed studies on a particular relation of interest. We also discussed that although reviewers and editors may contribute to publication bias, authors are likely the primary cause. Next, we described evidence showing that the majority of meta-analytic studies in the organizational sciences do not assess the potential presence of

TABLE 2.1 Recommendations

Recommendation	*Description*
Myth #1: Publication bias is concerned with the availability of all possible effect sizes in all areas of a scientific field.	
• Focus on specific relations of interest	– Investigations into the existence and prevalence of publication bias should focus on specific relations of interest.
Myth #2: The editorial review process is the primary cause of publication bias.	
• Disseminate results	– Authors should make efforts to disseminate their results regardless of the outcomes of their analyses.
Myth #3: There is a perception that the failsafe N and subgroup comparison are the best publication bias detection techniques.	
• Triangulate the mean effect size	– Meta-analytic researchers should address potential publication bias by using multiple advanced techniques (e.g., trim and fill, selection models, cumulative meta-analysis) in order to triangulate the meta-analytic mean effect size estimate.
• Test within subgroups	– Publication bias tests should be employed within homogeneous subgroups when feasible to decrease the possibility that heterogeneity distorts conclusions.
• Consider the influence of heterogeneity	– Future research should consider the influence of heterogeneity due to artifactual variance (e.g., measurement error), outliers, and moderators on the accuracy of publication bias tests.
Myth #4: Publication bias may not be of concern in the social sciences because it is prevented by reporting correlation matrices, testing multiple hypotheses, and conducting systematic searches in meta-analytic reviews.	
• Honor code	– Honor codes should be employed by journals to enhance the probability that submitting authors did not engage in any questionable research practices that might lead to outcome-level publication bias.
• Supplemental information online	– Supplemental information should be provided online by journals so that authors may provide additional information about their study design and the population of their study, as well as to report additional outcomes.
• Data sharing	– Journals should require that authors submit their raw data and syntax along with their submitted studies in order to increase transparency and to allow for replications. Journals could make the data available on their websites after a grace period (e.g., 3 years).
• Two-stage review process	– A two-stage review process should be implemented in which reviewers are only allowed to see the Results and Discussion sections of a study after they have reviewed the Introduction and Methods sections of the study.
• Incentives for B- and C-tier journals	– Universities should offer incentives for researchers to publish in B- and C-tier journals to keep them from abandoning studies when it appears no longer likely that they will be published in an A-tier journal.

(*Continued*)

TABLE 2.1 (Continued)

Recommendation	*Description*
• Journal submission and communications	– Journals should release original manuscript submissions as well as the communications among action editors, reviewers, and authors in order to provide greater transparency meant to reduce questionable research practices that might occur in the review process.
• Replications and prospective meta-analyses	– Journals should dedicate space solely for replication studies. Additionally, prospective meta-analyses should be implemented as means to encourage simultaneous replication studies.
• Study and protocol registration	– Registries should be created that allow for protocol registration prior to conducting a study so as to discourage questionable research practices and outcome-level publication bias that could occur after a study was conducted. Additionally, registries should be created that provide a repository of completed studies for meta-analytic researchers.

publication bias (Banks, Kepes, & McDaniel, 2012; Kepes et al., 2012). Furthermore, when meta-analytic studies do test for this bias, they typically use suboptimal methods. Finally, we discussed why the organizational sciences are not likely to be immune from the effects of publication bias. Specifically, reporting correlation matrices, testing multiple hypotheses, and conducting thorough systematic searches in meta-analytic reviews are unlikely to completely prevent the occurrence of publication bias. Many other prevention techniques should be used to mitigate the threat. By better educating researchers on the definition, causes, detection, and prevention of this bias, as a field we should be more able to assess and mitigate its effects.

Note

1 Consistent with the publication bias literature, we use the word "suppression" to refer to studies that are not published or otherwise readily available (Kepes & McDaniel, 2013). This use of the word "suppression" does not imply deceit.

References

American Psychological Association. (2008). Reporting standards for research in psychology: Why do we need them? What might they be? *American Psychologist, 63*, 839–851.

American Psychological Association. (2010). *Publication manual of the American Psychological Association (6th ed.).* Washington, DC: American Psychological Association.

Ariely, D. (2012). *The (honest) truth about dishonesty: How we lie to everyone—especially ourselves.* New York, NY: HarperCollins.

Banks, G. C., Batchelor, J. H., & McDaniel, M. A. (2010). Smarter people are (a bit) more symmetrical: A meta-analysis of the relationship between intelligence and fluctuating asymmetry. *Intelligence, 38*, 393–401.

Banks, G. C., Batchelor, J. H., Seers, A., O'Boyle, E. H., Pollack, J. M., & Gower, K. (2014). What does team-member exchange bring to the party? A meta-analytic review of team and leader social exchange. *Journal of Organizational Behavior, 35*, 273–295.

Banks, G. C., Kepes, S., & Banks, K. P. (2012). Publication bias: The antagonist of meta-analytic reviews and effective policy making. *Educational Evaluation and Policy Analysis, 34*, 259–277.

Banks, G. C., Kepes, S., & McDaniel, M. A. (2012). Publication bias: A call for improved meta-analytic practice in the organizational sciences. *International Journal of Selection and Assessment, 20*, 182–196.

Banks, G. C., & McDaniel, M. A. (2011). The kryptonite of evidence-based I-O psychology. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 4*, 40–44.

Banks, G. C., & O'Boyle, E. H. (2013). Why we need I-O psychology to fix I-O psychology. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 6*, 291–294.

Becker, B. J. (1994). Combining significance levels. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 215–230). New York, NY: Russell Sage Foundation.

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.

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics, 50*, 1088–1101.

Bennett, A. A., & Miao, C. (2013). How do we know truth? Extensions and examples from similar academic fields. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 6*, 276–278.

Berlin, J. A., & Ghersi, D. (2005). Preventing publication bias: Registries and prospective meta-analysis. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta analysis: Prevention, assessment, and adjustments* (pp. 35–48). West Sussex, UK: Wiley.

Biemann, T. (2013). What if we were Texas sharpshooters? Predictor reporting bias in regression analysis. *Organizational Research Methods, 16*, 335–363.

Blackwell, S. C., Thompson, L., & Refuerzo, J. (2009). Full publication of clinical trials presented at a national maternal-fetal medicine meeting: Is there a publication bias? *American Journal of Perinatology, 26*, 679–682.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* West Sussex, UK: Wiley.

Bosco, F. A., Field, J. G., & Pierce, C. A. (2012, August). *Accomodational plasticity in organizational science: Impact of hypothesis framing on effect size.* Paper presented at the annual meeting of the Academy of Management, Boston, MA.

Briner, R. B., & Rousseau, D. M. (2011). Evidence-based I-O psychology: Not there yet. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 4*, 3–22.

Chalmers, I., & Dickersin, K. (2013). Biased under-reporting of research reflects biased under-submission more than biased editorial rejection. *F1000Research, 2*, 1–6.

Chan, A.-W. (2008). Bias, spin, and misreporting: Time for full access to trial protocols and results. *PLoS Medicine, 5*, e230.

Chan, A.-W., & Altman, D. G. (2005). Identifying outcome reporting bias in randomised trials on PubMed: Review of publications and survey of authors. *British Medical Journal, 330*, 753–756.

Chiaburu, D. S., Peng, A. C., Oh, I.-S., Banks, G. C., & Lomeli, L. C. (2013). Employee organizational cynicism antecedents and outcomes: A meta-analysis. *Journal of Vocational Behavior, 83*, 181–197.

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.

Curfman, G. D., Morrissey, S., & Drazen, J. M. (2006). Expression of concern reaffirmed. *New England Journal of Medicine, 354*, 1193.

Dalton, D. R., Aguinis, H., Dalton, C. M., Bosco, F. A., & Pierce, C. A. (2012). Revisiting the file drawer problem in a meta-analysis: An assessment of published and nonpublished correlation matrices. *Personnel Psychology, 65*, 221–249.

De Angelis, C., Drazen, J. M., Frizelle, F. A., Haung, C., Hoey, J., Horton, R., Kotzin, S. . . . Van Der Weyden, M. B. (2004). Clinical trial registration: A statement from the International Committee of Medical Journal Editors. *New England Journal of Medicine, 351*, 1250–1251.

Dickersin, 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.

Duval, S. J., & Tweedie, R. L. (2000a). A nonparametric "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*, 455–463.

Dwan, K., Altman, D. G., Arnaiz, J. A., Bloom, J., Chan, A.-W., Cronin, E., Decullier, E. . . . Williamson, P. R. (2008). Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PloS One, 3*, e3081.

Easterbrook, P. J., Gopalan, R., Berlin, J. A., & Matthews, D. R. (1991). Publication bias in clinical research. *The Lancet, 337*, 867–872.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*, 629–634.

Emerson, G. B., Warme, W. J., Wolf, F. M., Heckman, J., Brand, R. A., & Leopold, S. S. (2010). Testing for the presence of positive-outcome bias in peer review: A randomized controlled trial. *Archives of Internal Medicine, 170*, 1934–1939.

Epstein, W. M. (1990). Confirmational response bias among social work journals. *Science, Technology, and Human Values, 15*, 9–37.

Etter, J. F., & Stapleton, J. (2009). Citations to trials of nicotine replacement therapy were biased toward positive results and high-impact-factor journals. *Journal of Clinical Epidemiology, 62*, 831–837.

Evans, S. (1996). Statistician's comment (to Misleading meta-analysis: "Fail safe N" is a useful mathematical measure of the stability of results by R. Persaud). *British Medical Journal, 312*, 125.

Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics, 90*, 891–904.

Ferguson, C. J., & Brannick, M. T. (2012). Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychological Methods, 17*, 120–128.

Gerber, A., & Malhotra, N. (2008a). Do statistical reporting standards affect what is published? Publication bias in two leading political science journals. *Quarterly Journal of Political Science, 3*, 313–326.

Gerber, A. S., & Malhotra, N. (2008b). Publication bias in empirical sociological research: Do arbitrary significance levels distort published results? *Sociological Methods & Research, 37*, 3–30.

Greenhouse, J. B., & Iyengar, S. (2009). Sensitivity analysis and diagnostics. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 417–433). New York, NY: Russell Sage Foundation.

Greenwald, A. G. (1975). Consequences of prejudice against the null hypothesis. *Psychological Bulletin, 82*, 1–20.

Hahn, S., Williamson, P. R., Hutton, J. L., Garner, P., & Flynn, E. V. (2000). Assessing the potential for bias in meta-analysis due to selective reporting of subgroup analyses within studies. *Statistics In Medicine, 19*, 3325–3336.

Harrison, J. S., Banks, G. C., Pollack, J. M., O'Boyle Jr., E. H., & Short, J. C. (in press). Publication bias in strategic management research. *Journal of Management.*

Hartshorne, J., & Schachner, A. (2012). Tracking replicability as a method of post-publication open evaluation. *Frontiers in Computational Neuroscience, 6*, 8.

Hedges, L. V., & Vevea, J. L. (2005). Selection methods approaches. In H. R. Rothstein, A. Sutton, & M. Borenstein (Eds.), *Publication bias in meta analysis: Prevention, assessment, and adjustments* (pp. 145–174). West Sussex, UK: Wiley.

Higgins, J. P., & Green, S. (Eds.). (2011). *Cochrane handbook for systematic reviews of interventions; Version 5.1.0 [updated March 2011]*: The Cochrane Collaboration. Available from www.cochrane-handbook.org

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, UK: Wiley.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.

Ioannidis, J. (2008). Interpretation of tests of heterogeneity and bias in meta-analysis. *Journal of Evaluation in Clinical Practice, 14*, 951–957.

Ioannidis, J., & Doucouliagos, H. (2013). What's to know about the credibility of empirical economics? *Journal of Economic Surveys, 27*, 997–1004.

Jick, T. D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly, 24*, 602–611.

Kepes, S., Banks, G. C., McDaniel, M. A., & Whetzel, D. L. (2012). Publication bias in the organizational sciences. *Organizational Research Methods, 15*, 624–662.

Kepes, S., Banks, G. C., & Oh, I. S. (2014). Avoiding bias in publication bias research: The value of "null" findings. *Journal of Business and Psychology*, 29, 183–203.

Kepes, S., Bennett, A., & McDaniel, M. A. (2014). Evidence-based management and the trustworthiness of our cumulative scientific knowledge: Implications for research, teaching, and practice. *Academy of Management Learning & Education, 13*, 446–466.

Kepes, S., & McDaniel, M. A. (2013). How trustworthy is the scientific literature in I-O psychology? *Industrial and Organizational Psychology: Perspectives on Science and Practice, 6*, 252–268.

Kepes, S., McDaniel, M. A., Banks, G. C., Hurtz, G. M., & Donovan, J. J. (2011, April). *Big five validity and publication bias: Conscientiousness' validity is lower than assumed*. Paper presented at the annual meeting of the Society for Industrial and Organizational Psychology, Chicago, IL.

Kepes, S., McDaniel, M. A., Brannick, M. T., & Banks, G. C. (2013). Meta-analytic reviews in the organizational sciences: Two meta-analytic schools on the way to MARS. *Journal of Business and Psychology, 28*, 123–143.

Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review, 2*, 196–217.

Kicinski, M. (2013). Publication bias in recent meta-analyses. *PLoS One, 8*, e81823.

Kirkham, J. J., Riley, R. D., & Williams, P. R. (2011). A multivariate meta-analysis approach for reducing the impact of outcome reporting bias in systematic reviews. *Statistics in Medicine, 31*, 2179–2195.

Leavitt, K., Mitchell, T. R., & Peterson, J. (2010). Theory pruning: Strategies to reduce our dense theoretical landscape. *Organizational Research Methods, 13*, 644–667.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48*, 1181–1209.

Mahoney, M. J. (1977). Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive Therapy & Research, 1*, 161–175.

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research, 45*, 633–644.

McDaniel, M. A. (2009, April). *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.

McDaniel, M. A., & Kepes, S. (in press). An evaluation of Spearman's hypothesis by manipulating g-saturation. *International Journal of Selection and Assessment.*

McDaniel, M. A., McKay, P., & Rothstein, H. (2006, April). *Publication bias and racial effects on job performance: The elephant in the room*. Paper presented at the annual meeting of the Society for Industrial and Organizational Psychology, Dallas, TX.

McDaniel, M. A., Rothstein, H. R., & Whetzel, D. L. (2006). Publication bias: A case study of four test vendors. *Personnel Psychology, 59*, 927–953.

McGauran, N., Wieseler, B., Kreis, J., Schuler, Y. B., Kolsch, H., & Kaiser, T. (2010). Review reporting bias in medical research—a narrative. *Trials, 11*, 1–15.

McKay, P. F., & McDaniel, M. A. (2006). A reexamination of black–white mean differences in work performance: More data, more moderators. *Journal of Applied Psychology, 91*, 538–554.

Murtaugh, P. A. (2002). Journal quality, effect size, and publication bias in meta-analysis. *Ecology, 83*, 1162–1166.

O'Boyle, E. H., Banks, G. C., & Gonzalez-Mule, E. (in press). The chrysalis effect: How ugly data metamorphosize into beautiful articles. *Journal of Management.*

O'Boyle, E. H., Rutherford, M., & Banks, G. C. (in press). Publication bias in entrepreneurship research: An examination of dominant relations to performance. *Journal of Business Venturing.*

Olson, C. M., et al. (2002). Publication bias in editorial decision making. *Journal of the American Medical Association, 287*, 2825–2828.

Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics, 8*, 157–159.

Palmer, T. M., Peters, J. L., Sutton, A. J., & Moreno, S. G. (2008). Contour-enhanced funnel plots for meta-analysis. *Stata Journal, 8*, 242–254.

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology, 61*, 991–996.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*, 638–641.

Rothstein, H. (2012). Accessing relevant literature. In H. M. Cooper (Ed.), *APA handbook of research methods in psychology: Vol. 1. Foundations, planning, measures, and psychometrics* (pp. 133–144). Washington, DC: American Psychological Association.

Rothstein, H. R., & Hopewell, S. (2009). Grey literature. In H. M. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 103–126). New York, NY: Russell Sage Foundation.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005). 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, UK: Wiley.

Sackett, P. R., & Larson, J. R. (1990). Research strategies and tactics in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 1, pp. 419–489). Palo Alto, CA: Consulting Psychologists Press.

Schmidt, F., & Oh, I. S. (2013). Methods for second-order meta-analysis and illustrative applications. *Organizational Behavior and Human Decision Processes, 121*, 204–218.

Schmidt, F. L., & Hunter, J. E. (2014). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Newbury Park, CA: Sage.

Schmidt, F. L., Hunter, J. E., Pearlman, K., Hirsh, H. R. (1985). Forty questions about validity generalization and meta-analysis. *Personnel Psychology, 38*, 697–798.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file drawer. *Journal of Experimental Psychology: General, 143*, 534–547.

Song, F., et al. (2010). Dissemination and publication of research findings: An updated review of related biases. *Health Technology Assessment, 14*, 1–220.

Sterling, T. D., & Rosenbaum, W. L. (1995). Publication decisions revisited: The effect of the outcome of statistical tests on the decision to publish and vice versa. *American Statistician, 49*, 108–112.

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, UK: Wiley.

Sterne, J.A.C., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J., Carpenter, J. . . . Higgins, J. P. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal, 343*, d4002.

Strech, D. (2011). The ethics of a restrictive regulation of trial registration. *Ethik in Der Medizin, 23*, 177–189.

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., pp. 435–452). New York, NY: Russell Sage Foundation.

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.

Tate, B. W., & McDaniel, M. A. (2008, August). *Race differences in personality: An evaluation of moderators and publication bias.* Paper presented at the annual meeting of the Academy of Management, Anaheim, CA.

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.

Turner, E. H., Matthews, A. M., Linardatos, E., Tell, R. A., & Rosenthal, R. (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine, 358*, 252–260.

Vevea, J. L., Clements, N. C., & Hedges, L. V. (1993). Assessing the effects of selection bias on validity data for the General Aptitude Test Battery. *Journal of Applied Psychology, 78*, 981–987.

Vevea, J. L., & Woods, C. M. (2005). Publication bias in research synthesis: Sensitivity analysis using a priori weight functions. *Psychological Methods, 10*, 428–443.

Weinhandl, E. D., & Duval, S. (2012). Generalization of trim and fill for application in meta-regression. *Research Synthesis Methods, 3*, 51–67.

Wertil, M. M., Schob, M., Brunner, F., & Steurer, J. (2013). Incomplete reporting of baseline characteristics in clinical trials: An analysis of randomized controlled trials and systematic reviews involving patients with chronic low back pain. *PLos One, 8*, e58512.

Whetzel, D. L. (2006, April). *Publication bias of studies on the validity of customer service measures: A review of Frei and McDaniel's (1998) Meta-analytic results.* Paper presented at the annual meeting of the Society for Industrial and Organizational Psychology, Dallas, TX.

Williamson, P. R., & Gamble, C. (2007). Application and investigation of a bound for outcome reporting bias. *Trials, 8*, 9.

Zarin, D. A., Tse, T., & Ide, N. C. (2005). Trial registration at ClinicalTrials.gov between May and October 2005. *New England Journal of Medicine, 353*, 2779–2787.

Zarin, D. A., Tse, T., Williams, R. J., Califf, R. M., & Ide, N. C. (2011). The ClinicalTrials.gov results database: Update and key issues. *New England Journal of Medicine, 364*, 852–860.