2013

The Science and Practice of Research Synthesis

Julia H. Littell
Bryn Mawr College, jlittell@brynmawr.edu

Journal of the Society for Social Work and Research
Volume 4, Issue 4, 292–299

December 31, 2013
ISSN 1948-822X DOI:10.5243/jsswr.2013.19

Guest Editor's Introduction to Special Issue:

The Science and Practice of Research Synthesis

Julia H. Littell
Bryn Mawr College

It is time to re-examine the methods that are commonly used to cumulate empirical evidence. Scientific methods for identifying, analyzing, and synthesizing results of multiple studies are rapidly developing, yet these methods are widely misunderstood and underutilized in the social, behavioral, and health sciences. In other words, the practice of research synthesis has not kept up with the science of research synthesis. Readers expect research reviews to provide comprehensive and accurate summaries of relevant bodies of evidence, but most published reviews fall far short of this goal (Gibbs, 2003; Bastian, Glasziou, & Chalmers, 2010). This special issue of the *Journal of the Society for Social Work and Research* focuses on systematic reviews and meta-analysis, which can be used in tandem to provide more accurate and informative syntheses of empirical research.

Research reviews are a staple of the professional and scholarly literature, and reviews can serve important functions. Because empirical knowledge is cumulative, careful syntheses of the results of multiple studies are essential for advances in theory and practice. Rigorous research synthesis can help practitioners and researchers understand overall trends and variations in the empirical literature, and resolve or explain differences that appear across reports that come from different studies, samples, and settings. Not only can reviews provide summaries of what is known, but (as explained below) systematic review methods can be used to develop hypotheses and answer questions not even considered in the original studies. Equally important, effective and accurate reviews can help decision makers cope with information overload.

The production of primary research reports began to rise in the 1960s and has not yet begun to plateau (Bastian et al., 2010). There are now more than 23 million citations in PubMed and 3.5 million records in PsycINFO. Most practitioners and other decision makers cannot keep up with the increasing volume of research, and many must rely on others to critically appraise and synthesize research results for them. As a result, research reviews tend to be cited more often than primary research reports. The higher citation count serves as an incentive for academics to produce research reviews.

The number of published research reviews increased dramatically in recent decades, first following and then far surpassing the production of primary research reports (Bastain et al., 2010). Although the number of new systematic reviews has increased steadily since the creation of the Cochrane Collaboration in the early 1990s, this trend has been dwarfed by the rapid proliferation of nonsystematic reviews in the scholarly literature. More than 75,000 nonsystematic reviews were published in 2007 alone, compared with fewer than 6,000 systematic reviews (and less than 25,000 trials) in the same year (Bastian et al., 2010).

Empirical evidence shows that nonsystematic reviews are affected by many sources and types of bias, and can lead to the wrong conclusions (Bushman & Wells, 2001; Carlton & Strawderman, 1996; Cooper & Rosenthal, 1980; Littell, 2008; Mann, 1994). Yet, despite the great potential for error in nonsystematic reviews (Goldschmidt, 1986; Mulrow, 1987), even in medicine, "the staple of…literature synthesis remains the nonsystematic narrative review" (Bastian et al., 2010, p. 4).

Twenty years ago, Sir Iain Chalmers and his colleagues noted most research scientists

> … *operate on a double standard: they go to great lengths to define the methods they used to minimize biases and random errors in their reports on the results of new research, but they often do not attempt to apply scientific principles in their discussions of how the newly generated evidence accords with previously available information. Scientists also operate by this double standard when they conduct and report…[research] reviews* (Chalmers, Enkin, & Keirse, 1993, p. 411-412).

These sentiments were echoed by Mark Lipsey in 1997, when he likened evaluation research to the construction of bricks. Lipsey noted that evaluators were consumed with the properties of primary studies, but paid far less attention to methods for building knowledge and theory from dozens or thousands of studies. Meta-analysis is an important tool for constructing theory and scientific evidence, but meta-analysis alone is insufficient.

The scientific approach to research synthesis, as articulated by Harris Cooper (1982) and others, treats

the entire review process as a form of scientific inquiry. Reviews are observational investigations, akin to surveys, in which the unit of analysis is the study (or nonoverlapping sample). Basic principles of science apply to research reviews as well as to primary studies: if we care about the validity of inferences drawn from empirical data, then our study designs, sampling procedures, data collection methods, and analytic techniques matter, regardless of whether we are studying individuals, communities, or previous studies.

Too often, basic principles of science fall by the wayside when scholars conduct research reviews. Methods and inferences that are generally not accepted in primary research—including generalizations based on convenience samples, unreliable data collection procedures, outmoded analyses, and anecdotal reports—routinely appear in published reviews. Indeed, traditional nonsystematic reviews tend to rely on convenience samples of published studies. Inclusion criteria and other decision rules are rarely clear, so that readers might not know why some studies were included and others were excluded. Justifications for haphazard review methods include ad hominem arguments (reviewers claim special expertise on the topic) and appeals to tradition or authority (reviewers followed procedures used in the past). These explanations are inadequate, given the extensive body of empirical literature on bias and error in research reviews, and review methods that minimize these problems.

Sources of Bias and Error in Reviews

Sources of bias and error in research reviews are well known. These include problems that arise (a) in the original studies, (b) in the reporting and dissemination of research results, and (c) in the review process itself (Littell, 2008). As Chalmers and colleagues (1993) pointed out, much attention has been paid to methods for limiting bias and error in primary studies. Below I consider the last two sources of error.

When reviewers use key word searches of electronic databases, they obtain convenience samples of studies. These samples are not representative of all studies conducted on the topic. Rather, these samples are likely to be biased by selective reporting, publication, and dissemination patterns that favor statistically significant, positive effects (Dwan et al., 2008; Hopewell, Loudon, Clarke, Oxman, & Dickersin, 2009; Song, Parekh, et al., 2010; Song, Parekh-Bhurke, et al., 2009). Indeed, empirical evidence shows that:

- incomplete and selective reporting of research results is common (Dwan et al., 2008; Pigott, Valentine, Polanin, Williams, & Canada, 2013; Smyth et al., 2011);
- fewer than half of all completed studies are published (Chalmers, Glasziou, & Godlee, 2013; Cooper, DeNeve, & Charlton, 1997; Dwan et al., 2008; Jones et al., 2013);
- publication status is not a good proxy for study quality (McLeod & Weisz, 2004); and
- the odds of publication are 2 to 3 times greater for studies that have positive, statistically significant results (Song, Parekh, et al., 2010; Song, Parekh-Bhurke, et al., 2009).

Further, studies with statistically significant results are published more quickly (Hopewell, Clarke, Stewart, & Tierney, 2007) and cited and reprinted more often (Egger & Smith, 1998) than other studies. These reporting, publication, and dissemination biases tend to inflate some effect sizes, perpetuate unfounded beliefs, and distort research results (Greenberg, 2009; Young, Ioannidis, & Al-Ubaydli, 2008). Thus, reviewers must be careful to obtain unbiased samples of studies.

Data extraction errors are common in reviews (Ford, Guyatt, Talley, & Moayyedi, 2010; Gøtzsche, Hrógjartsson, Maric, & Tendal, 2007). Initial agreement between trained observers is low (Tendal et al., 2009), but experimental evidence shows data extraction by a single reviewer results in more error than duplicate extraction and resolution of initial differences (Buscemi, Hartling, Vandermeer, Tjosvold, & Klassen, 2005).

Experimental evidence shows that narrative synthesis is less accurate than meta-analysis (Bushman & Wells, 2001; Cooper & Rosenthal, 1980). A statistical technique known as *vote counting* is sometimes applied in nonsystematic reviews. Vote counting relies on the statistical significance of the outcomes in the primary studies. When the primary studies are not carried out with high levels of statistical power — which is often the case in research in the behavioral and social sciences — vote counting has the undesirable property of having *less* statistical power as *more* evidence cumulates (Hedges & Olkin, 1985). Further, voting counting can easily lead to wrong conclusions because between-study variations in sample size and heterogeneity are usually not taken into account (Carlton & Strawderman, 1996; Mann, 1994). Informal versions of vote counting appear in some nonsystematic reviews in statements about results of "most studies" or the "weight of the evidence." These statements appear to be based on cognitive algebra, which is often inaccurate (Valentine & Cooper, 2008).

Systematic Review Methods

Systematic reviews aim to minimize bias and error at each step in the review process. When paired with meta-analysis and related techniques, these methods are well equipped to handle many synthesis problems. Systematic reviews are not limited to questions about intervention effects, nor are they limited to randomized controlled trials.

Results of methodological research on strategies to reduce bias and error in reviews have been used to create evidence-based guidelines for the conduct and reporting of systematic reviews and meta-analyses. These guidelines are available from the Cochrane Collaboration (Chandler, Churchill, Higgins, Lasserson, & Tovey, 2013; Higgins & Green, 2011), the Institute of Medicine (IOM; 2011), and the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) group (Moher, Liberati, Tetzlaff, Altman, & the PRISMA Group, 2009; Welch et al., 2012). In addition, the AMSTAR tool Assessment of Multiple SysTemAtic Reviews) was developed for rapid assessment of systematic reviews (Shea et al., 2007).

To increase transparency and limit opportunities for bias, systematic reviewers develop and follow a predetermined protocol for the review. Protocols are made public in advance of the review (Stewart, Moher, & Shekelle, 2012). Reviewers use transparent (i.e., well-documented and replicable) procedures to locate, analyze, and synthesize results of previous studies.

Logic models can be used to clarify the scope, central questions, and hypotheses for a review (Anderson et al., 2011). Formal eligibility criteria for systematic reviews are developed in advanced, often using the PICOS framework (populations, interventions, comparisons, outcomes, and study designs; Higgins & Green, 2011).

The search for eligible studies is treated as a sampling problem: systematic reviews aim for a representative sample of all studies that meet eligibility criteria. Reviewers collaborate with search specialists and librarians to develop sensitive and specific search strategies, proper application of search filters, and documentation of the search process in sufficient detail for replication (Hammerstrøm, Wade, & Jørgensen, 2010). Grey literature searches are usually included to reduce the risk of publication bias (Hopewell, McDonald, Clarke, & Egger, 2007).

Study eligibility decisions are made by two or more reviewers, who work independently and then compare results. Specific reasons for study exclusion are documented. Important study characteristics and results are extracted onto structured coding sheets. Emphasis is on achieving high inter-rater reliability; hence, double extraction is generally required (Chandler et al., 2013; IOM, 2011).

The Cochrane risk of bias framework is often used for assessing randomized controlled trials (Higgins & Green, 2011); extensions of this framework for nonrandomized studies are under development. Overall study quality scores have been shown to have questionable reliability and validity (Jüni, Altman, & Egger, 2001; Valentine & Cooper, 2008) because such scores conflate unrelated methodological issues and study design or implementation features, which might have different impacts on reliability or validity. Emphasis is on examining the separate influence of key components of methodological quality (Wells & Littell, 2009). The GRADE system (Grades of Recommendation Assessment, Development and Evaluation) is often used to evaluate the quality of evidence across studies; for each relevant outcome in a systematic review, GRADE assesses the amount of data available, consistency of results, overall risk of bias, and confounding variables across studies (Guyatt, Oxman, Schünemann, Tugwell, & Knotterus, 2010).

Meta-analysis can be used if two or more studies provide quantitative data on the same variable. Random effects models are typically used in reviews of complex psychosocial interventions and other topics in which heterogeneous results are expected (Bornstein, Hedges, Higgins, & Rothstein, 2009). In larger reviews, moderator analysis can be used to identify characteristics of participants, treatments, study designs, or contexts that are associated with variations in results. Multivariate methods have been developed to assess intervention effects on multiple dependent variables (Hedges, Tipton, & Johnson, 2010; Tanner-Smith & Tipton, 2013). Network meta-analysis is a relatively new technique that provides direct and indirect comparisons of multiple interventions for a single condition; network meta-analysis can be used to rank the effectiveness of interventions that have never been compared in head-to-head trials (Salanti, 2012).

Methods to detect and adjust for publication bias in meta-analysis are usually considered. The Failsafe N (or file drawer number) has been abandoned by meta-analysts in favor of more robust techniques to assess and correct for publication bias and related problems. Funnel plots and contour-enhanced funnel plots are commonly used, along with formal methods for assessment of funnel plot asymmetry (such as trim-and-fill analysis and Egger's test; see Rothstein, Sutton, & Bornstein, 2005). Newer regression methods are also available (Moreno et al., 2009).

Reviewers often translate results of systematic reviews and meta-analysis into metrics that are more

meaningful to policy makers and practitioners (Lipsey & Wilson, 2001). For example, Wilson and Tanner-Smith (2013) illustrate how odds ratios can be converted into more practical indicators of program impact.

A variety of aggregative and interpretive methods have been used to synthesize qualitative evidence, either alone or to complement quantitative synthesis. For example, synthesis of qualitative evidence might enrich understanding of consumers' experiences, mechanisms of change, or adverse effects of intervention. However, methods for quantitative synthesis are under debate and there is little empirical evidence of the robustness of various qualitative synthesis methods (Noyes, Popay, Pearson, Hannes, & Booth, 2011).

Misuse of the term "systematic review." With growing interest in systematic reviews, misconceptions have flourished and the term *systematic review* has become widely misused. This term has been incorrectly applied to reviews that were based solely on convenience samples of published studies, made no effort to ensure reliable data extraction, and used narrative synthesis or vote-counting when better methods were available.

Underutilization and misuse of scientific methods of research synthesis can be attributed, in part, to the dearth of formal training in these methodologies in doctoral and post-doctoral programs in the behavioral, social, and health sciences. Another plausible explanation is that scholars are reluctant to give up easy, familiar practices for new and more difficult ones, even when empirical evidence shows that the latter are more reliable and accurate. Ironically, nonsystematic reviews are often used to promote "evidence-based practices," even though these reviews are not based on the best available evidence about how to review research. Consistent with the concept of confirmation bias, it is easier to suggest that *others* seek and use new evidence to inform *their* decisions than to do this ourselves.

Scientific research syntheses are essential for building a reliable base of empirical evidence. Thus, in 2013, the *Journal of the Society for Social Work and Research* adopted the policy that authors of reports on systematic reviews and meta-analyses should follow evidence-based guidelines for the conduct and reporting of systematic reviews and meta-analyses.

In This Issue

This special issue of the *Journal of the Society for Social Work and Research* includes four reports on systematic reviews and two papers on research synthesis methodologies. These papers illustrate the groundbreaking work of the Cochrane and Campbell Collaborations and of members of the interdisciplinary Society for Research Synthesis Methodology.

Bennett, Barlow, Huband, Smailagic, and Roloff describe their Cochrane review of 48 randomized controlled trials on effects of group parenting programs on psychosocial outcomes for parents. These programs produced short-term improvements in parents' well-being, but improvements were not maintained at one-year follow-ups.

Pitt, Lowe, Prictor, Hetrick, Ryan, Berends, and Hill report on their Cochrane review of research on consumer-providers' effects on client outcomes in mental health services. Five randomized controlled trials compared services provided by consumer-providers with those provided by professionals employed to perform the same roles. Six trials compared effects of professional mental health services with and without the aid of consumer-providers. No differences in outcomes were found, whether consumer-providers substituted for or added to services provided by professionals. Authors discuss ways in which evidence in this area could be strengthened.

Wilson and Tanner-Smith describe their Campbell Collaboration review of 152 studies of school drop out prevention and intervention programs. This review includes both randomized and quasi-experimental designs. The authors made extensive efforts to find relevant unpublished studies and, consistent with the literature on publication bias, they found that relatively few (22%) of the studies on drop out programs were published in peer-reviewed journals. Randomized trials were more likely to appear in technical reports and dissertations than in peer-reviewed journal articles. Authors used robust variance estimates (Hedges et al., 2010; Tanner-Smith & Tipton, 2013) to assess program effects on multiple (correlated) effect sizes. Wilson and Tanner-Smith corrected for clustering effects and imputed missing data on moderator variables. The authors translated odds ratios into more accessible metrics (drop out rates). Moderator analyses showed that study methods and program implementation quality were associated with effect sizes, but effects appeared to be consistent across different types of programs and samples.

Steinka-Fry, Wilson, and Tanner-Smith analyzed a subset of studies from the previous review, focusing on effects of drop out prevention and intervention programs for pregnant and parenting teens. Their report includes 51 effect sizes from 15 studies. Overall, programs were effective in reducing school drop out and increasing school retention among pregnant and parenting adolescents. Weaker research designs and greater implementation quality were associated with larger effects.

Aloe and Thompson discuss the synthesis of partial effect sizes from studies that report regression models and in situations when it is important to partial out effects of certain variables. This article illustrates uses of the correlation family of effect sizes meta-analysis with observational data beyond questions about intervention effects. One of the remaining challenges for reviewers is determining how to handle partial effect sizes derived from regression models that included different covariates.

Grant and Calderbank-Batista provide an introduction to network meta-analysis (NMA). Hailed as "the next generation evidence synthesis tool" (Salanti, 2012, p. 80), NMA uses direct and indirect evidence to rank the effectiveness of alternative interventions for a specific condition. Although NMA is considered the best available technique for ranking the effectiveness of alternative treatments, NMA has rarely been used outside of medicine. Grant and Calderbank-Batista provide a cogent discussion of underlying assumptions of NMA (e.g., transitivity and consistency) and identify special considerations for using this approach in reviews of complex psychosocial interventions.

Together, these articles provide useful examples of the state of the science of research synthesis. They also offer intriguing glimpses into future work in these areas.

Future Directions

To build a reliable evidence base for practice and policy, we need more systematic reviews, better systematic reviews, more frequent updates of existing systematic reviews, and fewer nonsystematic reviews (Bastian et al., 2010). Chalmers and colleagues argued that systematic reviews should be conducted at the beginning and end of each new study to avert avoidable waste of research and related resources (Clarke, Hopewell, & Chalmers, 2010; Chalmers & Glasziou, 2009). Time and effort currently devoted to production of nonsystematic reviews could be greatly reduced or eliminated.

Recent studies point to the need for improvements in the conduct and reporting of published systematic reviews and meta-analyses (Aytug, Rothstein, Zhou, & Kern, 2012; Moher, Tetzlaff, Tricco, Sampson, & Altman, 2007). Because the growth in scholarly production has not yet begun to plateau, we will need leaner, more efficient systematic review methods to help decision makers keep up with the evidence (Bastian et al., 2010). For example, text-mining and machine-assisted screening of titles and abstracts will soon enable reviewers to quickly sort through thousands of studies with no loss of accuracy (Shemilt et al., 2013; Wallace, Trikalinos, Lau, Brodley, & Schmid, 2010). Efforts are also underway to improve access to unpublished studies (Chalmers et al., 2013). These and other advances on the horizon should help bridge current gaps between the science and practice of research synthesis.

More comprehensive, accurate, and up-to-date research syntheses can accelerate advancements in the social, behavioral, and health sciences and lead to better-informed, more effective efforts to improve public health and well-being.

Guest Editor Note

Julia H. Littell, PhD, is a professor in the Graduate School of Social Work and Social Research of Bryn Mawr College. She is co-editor-in-chief and co-chair of the Campbell Collaboration and a founding member of the Society for Research Synthesis Methodology.

Correspondence regarding the Guest Editor's Introduction to the special issue should be directed to Dr. Julia Littell, 300 Airdale Rd., Bryn Mawr, PA 19003 or via e-mail to jlittell@brynmawr.edu

References

Aloe, A. M., & Thompson, C. G. (2013). The synthesis of partial effect sizes. *Journal of the Society for Social Work and Social Research, 4*, 390-405. doi:10.5243/jsswr.2013.24

Anderson, L. M., Petticrew, M., Rehfuess, E., Armstrong, R., Ueffing, E., Baker, P., …Tugwell, P. (2011). Using logic models to capture complexity in systematic reviews. *Research Synthesis Methods, 2*, 33–42. doi:10.1002/jrsm.32

Aytug, Z. G., Rothstein, H. R., Zhou, W., & Kern, M. C. (2012). Revealed or concealed? Transparency of procedures, decisions, and judgment calls in meta-analyses. *Organizational Research Methods, 15*, 103–133. doi:10.1177/1094428111403495

Bastian, H., Glasziou, P., & Chalmers, I. (2010). Seventy-five trials and eleven systematic reviews a day: How will we ever keep up? *PLoS Med, 7*(9), e1000326. doi:10.1371/journal.pmed.1000326

Bennett, C., Barlow, J., Huband, N., Smailagic, N., & Roloff, V. (2013). Group-based parenting programs for improving parenting and psychosocial functioning: A systematic review. *Journal of the Society for Social Work and Research, 4*, 300-332. . doi:10.5243/jsswr.2013.20

Bornstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex, England: John Wiley & Sons. doi:10.1002/9780470743386

Buscemi, N., Hartling, L., Vandermeer, B., Tjosvold, L., & Klassen, T. P. (2005). Single data extraction generated more errors than double data extraction in systematic reviews. *Journal of Clinical Epidemiology*, *59*, 697–703. doi:10.1016/j.jclinepi.2005.11.010

Bushman, B. J., & Wells, G. L. (2001). Narrative impressions of literature: The availability bias and the corrective properties of meta-analytic approaches. *Personal and Social Psychology Bulletin*, *27*, 1123–1130. doi:10.1002/9780470743386

Carlton, P. L., & Strawderman, W. E. (1996). Evaluating cumulated research I: The inadequacy of traditional methods. *Biological Psychiatry*, *39*, 65–72. doi:10.1016/0006-3223(95)00089-5

Chalmers, I., Enkin, M., & Keirse, M. J. N. C. (1993). Preparing and updating systematic reviews of randomized controlled trials of health care. *Milbank Quarterly*, *71*(3), 411–437. doi:10.2307/3350409

Chalmers, I., & Glasziou, P. (2009). Avoidable waste in the production and reporting of research evidence. *Lancet*, *374*, 86–89. doi:10.1016/S0140-6736(09)60329-9

Chalmers, I., Glasziou, P., & Godlee, F. (2013). All trials must be registered and the results published: Academics and non-commercial funders are just as guilty as industry. *British Medical Journal*, *346*, f105. doi:10.1136/bmj.f105

Chandler, J., Churchill, R., Higgins, J., Lasserson, T., & Tovey, D. (2013). *Methodological standards for the conduct of new Cochrane Intervention Reviews (Version 2.3)*. Retrieved from http://www.editorial-unit.cochrane.org/mecir

Clarke, M., Hopewell, S., & Chalmers, I. (2010). Clinical trials should begin and end with systematic reviews of relevant evidence: 12 years and waiting. *Lancet*, *376*(9734), 20–21. doi:10.1016/S0140-6736(10)61045-8

Cooper, H. M. (1982). Scientific guidelines for conducting integrative research reviews. *Review of Educational Research*, *52*(2), 291–302. doi:10.3102/00346543052002291

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*(4), 447–452. doi:10.1037/1082-989X.2.4.447

Cooper, H. M., & Rosenthal, R. (1980). Statistical versus traditional procedures for summarizing research findings. *Psychological Bulletin*, *87*(3), 442–449. doi:10.1037/0033-2909.87.3.442

Dwan, K., Altman, D. G., Arnaiz, J. A., Bloom, J., Chan, A.-W., Cronin, E., … Williamson, P. R. (2008). Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS ONE*, *3*(8), e3081. doi:10.1371/journal.pone.0003081

Egger, M., & Smith, G. D. (1998). Bias in location and selection of studies. *British Medical Journal*, *316*(7124), 61–66. doi:10.1136/bmj.316.7124.61

Ford, A. C., Guyatt, G. H., Talley, N. J., & Moayyedi, I. P. (2010). Errors in the conduct of systematic reviews of pharmacological interventions for irritable bowel syndrome. *American Journal of Gastroenterology*, *105*(2), 280–288. doi:10.1038/ajg.2009.658

Gibbs, L. E. (2003). *Evidence-based practice for the helping professions: A practical guide with integrated multimedia*. Pacific Grove, CA: Brooks/Cole-Thompson Learning.

Goldschmidt, P. G. (1986). Information synthesis: A practical guide. *Health Services Research*, *21*(2), 215–237. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1068946/pdf/hsresearch00501-0091.pdf

Gøtzsche, P. C., Hrógjartsson, A., Maric, K., & Tendal, B. (2007). Data extraction errors in meta-analyses that use standardized mean differences. *JAMA*, *298*(4), 430–437. doi:10.1001/jama.298.4.430

Grant, E. S., & Calderbank-Batista, T. (2013). Network meta-analysis for complex social interventions: Problems and potential. *Journal of the Society for Social Work and Social Research*, *4*, 406–420. doi:10.5243/jsswr.2013.25

Greenberg, S. A. (2009). How citation distortions create unfounded authority: Analysis of a citation network. *British Medical Journal*, *339*. doi:10.1136/bmj.b2680

Guyatt, G. H., Oxman, A. D., Schünemann, H. J., Tugwell, P., & Knotterus, A. (2010). GRADE guidelines: A new series of articles in the Journal of Clinical Epidemiology. *Journal of Clinical Epidemiology*, *64*, 380-382. doi:10.1016/j.jclinepi.2010.09.011

Hammerstrøm, K., Wade, A., & Jørgensen, A.-M. K. (2010). *Searching for studies: A guide to information retrieval for Campbell systematic reviews*. Retrieved from

http://www.campbellcollaboration.org/resources/research/new_information_retrieval_guide.php

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*(1), 39–65. doi:10.1002/jrsm.5

Higgins, J. P. T., & Green, S. (Eds.). (2011). *Cochrane handbook for systematic reviews of interventions, Version 5.1.0.* Oxford, UK: The Cochrane Collaboration. Retrieved from http://handbook.cochrane.org

Hopewell, S., Clarke, M. J., Stewart, L., & Tierney, J. (2007). Time to publication for results of clinical trials. *Cochrane Database of Systematic Reviews*, *2007*(Issue 2), No. MR000011. doi:10.1002/14651858.MR000011.pub2

Hopewell, S., Loudon, K., Clarke, M. J., Oxman, A. D., & Dickersin, K. (2009). Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database of Systematic Reviews*, *2009, Issue 1*, Art. No. MR000006. doi:10.1002/14651858.MR000006.pub3

Hopewell, S., McDonald, S., Clarke, M., & Egger, M. (2007). Grey literature in meta-analyses of randomized trials of health care interventions. *The Cochrane Database of Systematic Reviews*, *2007*(Issue 2), No. MR000010. doi:10.1002/14651858.MR000010.pub3

Institute of Medicine. (2011). *Finding what works in health care: Standards for systematic reviews*. Washington, DC: National Academies Press. Retrieved from http://www.iom.edu/Reports/2011/Finding-What-Works-in-Health-Care-Standards-for-Systematic-Reviews/Standards.aspx

Jones, C. W., Handler, L., Crowell, K. E., Keill, L. G., Weaver, M. A., & Platts-Mills, T. F. (2013). Non-publication of large randomized clinical trials: Cross-sectional analysis. *British Medical Journal*, *347*, f6104. doi:10.1136/bmj.f6104

Jüni, P., Altman, D. G., & Egger, M. (2001). Assessing the quality of controlled clinical trials. *British Medical Journal*, *323*(7303), 42–46. doi:10.1136/bmj.323.7303.42

Lipsey, M. W. (1997). What can you build with thousands of bricks? Musings on the cumulation of knowledge in program evaluation. *New Directions for Evaluation*, 76, 7–23. doi:10.1002/ev.1084

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

Littell, J. H. (2008). Evidence-based or biased? The quality of published reviews of evidence-based practices. *Children and Youth Services Review*, *30*(11), 1299–1317. doi:10.1016/j.childyouth.2008.04.001

Mann, C. C. (1994). Can meta-analysis make policy? *Science*, *266*(5187), 960–962. doi:10.1126/science.7973676

McLeod, B. D., & Weisz, J. R. (2004). Using dissertations to examine potential bias in child and adolescent clinical trials. *Journal of Consulting and Clinical Psychology*, *72*(2), 235–251. doi:10.1037/0022-006X.72.2.235

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*, e1000097. doi:10.1371/journal.pmed.1000097

Moher, D., Tetzlaff, J., Tricco, A. C., Sampson, M., & Altman, D. G. (2007). Epidemiology and reporting characteristics of systematic reviews. *PLoS Med*, *4*(3), e78. doi:10.1371/journal.pmed.0040078

Moreno, S. G., Sutton, A. J., Ades, A. E., Stanley, T. D., Abrams, K. R., Peters, J. L., & Cooper, N. J. (2009). Assessment of regression-based methods to adjust for publication bias through comprehensive simulation study. *BMC Medical Research Methodology*, *9*(2). doi:10.1186/1471-2288-9-2

Mulrow, C. D. (1987). The medical review article: State of the science. *Annals of Internal Medicine*, *106*(3), 485–488. doi:10.7326/0003-4819-106-3-485

Noyes, J., Popay, J., Pearson, A., Hannes, K., & Booth, A. (2011). Qualitative research and Cochrane reviews. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions, Version 5.1.0.* Oxford, UK: The Cochrane Collaboration. Retrieved from http://handbook.cochrane.org

Pigott, T. D., Valentine, J. C., Polanin, J. R., Williams, R. T., & Canada, D. D. (2013). Outcome-reporting bias in education research. *Educational Researcher*, *43*, 424-432. doi:10.3102/0013189X13507104

Pitt, V. J., Lowe, D., Prictor, M., Hetrick, S., Ryan, R., Berends, L., & Hill, S. (2013). A systematic review of consumer-providers' effects on client outcomes in statutory mental health services: The evidence and the path beyond. *Journal of the Society for Social Work and Research, 4, 333–356.* doi:10.5243/jsswr.2013.21

Rothstein, H., Sutton, A. J., & Bornstein, M. (2005). *Publication bias in meta-analysis: Prevention, assessment, and adjustments*. Chichester, UK: John Wiley & Sons. doi:10.1002/0470870168

Salanti, G. (2012). Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: Many names, many benefits, many concerns for the next generation evidence synthesis tool. *Research Synthesis Methods*, *2012*(3), 80–97. doi:10.1002/jrsm.1037

Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., … Bouter, L. M. (2007). Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology*, *7*(10). doi:10.1186/1471-2288-7-10

Shemilt, I., Simon, A., Hollands, G. J., Marteau, T. M., Ogilvie, D., O'Mara-Eves, A., …Thomas, J. (2013). Pinpointing needles in giant haystacks: Use of text mining to reduce impractical screening workload in extremely large scoping reviews. *Research Synthesis Methods*. Advanced online publication. doi:10.1002/jrsm.1093

Smyth, R. M. D., Kirkham, J. J., Jacoby, A., Altman, D. G., Gamble, C., & Williamson, P. R. (2011). Frequency and reasons for outcome reporting bias in clinical trials: Interviews with trialists. *British Medical Journal, 342*. doi:10.1136/bmj.c7153

Song, F., Parekh, S., Hooper, L., Loke, Y. K., Ryder, J., Sutton, A. J.,Hing, C, Kwok, C. S., Pang, C., & Harvey, I. (2010). Dissemination and publication of research findings: An updated review of related biases. *Health Technology Assessment*, *14*(8), 1–220. doi:10.3310/hta14080 Retrieved from http://www.hta.ac.uk/fullmono/mon1408.pdf

Song, F., Parekh-Bhurke, S., Hooper, L., Loke, Y., Ryder, J., Sutton, A., & Harvey, I. (2009). Extent of publication bias in different categories of research cohorts: A meta-analysis of empirical studies. *BMC Medical Research Methodology*, *9*(1), 79. doi:10.1186/1471-2288-9-79

Steinka-Fry, K. T., Wilson, S. J., & Tanner-Smith, E. E. (2013). Effects of school drop out prevention programs for pregnant and parenting adolescents: A meta-analytic review. *Journal of the Society for Social Work and Research, 4*, 373–389. doi:10.5243/jsswr.2013.23

Stewart, L., Moher, D., & Shekelle, P. (2012). Why prospective registration of systematic reviews makes sense. *Systematic Reviews, 1*(7). doi:10.1186/2046-4053-1-7

Tanner-Smith, E. E., & Tipton, E. (2013). Robust variance estimation with dependent effect sizes: Practical considerations including a software tutorial in Stata and SPSS. *Research Synthesis Methods*. Advanced online publication. doi:10.1002/jrsm.1091

Tendal, B., Higgins, J. P. T., Jüni, P., Hróbjartsson, A., Trelle, S., Nüesch, E., … Gøtzsche, P. C. (2009). Disagreements in meta-analyses using outcomes measured on continuous or rating scales: Observer agreement study. *British Medical Journal, 339*, b3128. doi:10.1136/bmj.b3128

Valentine, J. C., & Cooper, H. (2008). A systematic and transparent approach for assessing the methodological quality of intervention effectiveness research: The Study Design and Implementation Assessment Device (Study DIAD). *Psychological Methods, 13*, 130-149. doi:10.1037/1082-989X.13.2.130

Wallace, B. C., Trikalinos, T. A., Lau, J., Brodley, C., & Schmid, C. H. (2010). Semi-automated screening of biomedical citations for systematic reviews. *BMC Bioinformatics*, *11*, 55. doi:10.1186/1471-2105-11-55

Wells, K., & Littell, J. H. (2009). Study quality assessment in systematic reviews of research on intervention effects. *Research on Social Work Practice, 19*, 52–62. doi:10.1177/1049731508317278

Welch, V., Petticrew, M., Tugwell, P., Moher, D., O'Neill, J., Waters, E., White, H., & the PRISMA-Equity Bellagio group. (2012). PRISMA-Equity 2012 extension: Reporting guidelines for systematic reviews with a focus on health equity. *PLoS Med*, *9*(10), 31001333. doi:10.1371/journal.pmed.1001333

Wilson, S. J., & Tanner-Smith, E. E. (2013). Drop out prevention and intervention programs for improving school completion among school-aged children and youth: A systematic review. *Journal of the Society for Social Work and Research, 4,* 357–372. doi:10.5243/jsswr.2013.22

Young, N. S., Ioannidis, J. P. A., & Al-Ubaydli, O. (2008). Why current publication practices may distort science. *PLoS Medicine*, *5*(10), e201. doi:10.1371/journal.pmed.0050201

Publication date: December 31, 2013