

## **Abstract Title Page**

**Title:** Examination of Biased Outcome Reporting in Educational Research

**Author(s):** Terri D. Pigott, Jeffrey C. Valentine, Ryan T. Williams, & Dericka D. Canada

## **Abstract Body**

### **Background/context:**

Publication bias is a term that typically refers to the well-known tendency for studies lacking statistically significant results to be less likely to be published in peer-reviewed journals. This happens because authors are less likely to submit, while editors and reviewers are less likely to accept for publication, papers that lack statistically significant results for their primary outcomes (see Dickersin, 2005 for a review). Some of the most notable cases of publication bias come from the medical sciences. Most recently, a high-profile examination of the impact of publication bias in anti-depressant research revealed that a startling 31% Phase II and Phase III trials have gone unpublished (Turner et al., 2008). The authors found that outcomes within studies (and some studies themselves) were more likely to be published if the results were positive (i.e., the drug is effective, and side effects appear to be minimal), while negative results were more likely not to be reported. Because published studies are easier to find than unpublished studies, later users of the studies, for example, scholars conducting a systematic review, can come away with a biased impression of the strength and nature of the evidence. Furthermore, as Sutton & Pigott (2005) point out, data censoring can manifest itself in a variety of ways. The possibilities for bias include missing whole studies (i.e., what is typically referred to as publication bias), missing outcomes (outcome reporting bias; Chan & Altman, 2005), missing subgroup analyses, and missing important study details (e.g., study participant information; details about the nature of the intervention).

### **Purpose / objective / research question / focus of study:**

Given the increased use of comprehensive literature reviews and meta-analysis for policy decisions, a thorough understanding of the prevalence and impact of data censoring in education

research is needed. The purpose of this investigation is to investigate systematically the prevalence and impact of publication bias in the educational sciences. In medicine, studies of outcome reporting bias have used registries of trials and their protocols to examine the reporting accuracy of subsequent publications of these registered trials. Since educational trials are not registered, we will use doctoral dissertations and their subsequent published reports to examine the prevalence and effects of outcome reporting bias in educational intervention research. We focus on intervention studies since the target outcomes are more easily identified than in observational studies. For example, a typical experimental or quasi-experimental study will identify a set of outcomes that will be compared across treatment groups whereas observational studies may have a range of potential goals such as the estimation of a multivariate model of sets of measures, or a structural model examining the interrelationships between measured constructs.

#### **Data Collection and Analysis:**

The first step in this study will involve a systematic search of the literature for dissertations that examine the effects of educational interventions (broadly construed). Starting with the Carnegie Foundation's list of 96 "very high" research productivity universities, we will use the database ProQuest Digital for relevant dissertations that were completed between the years of 2001 and 2005. There are two reasons for these specific date limits in the search. First, we are interested in inferences about the current state of the problem of data censoring so we intend to examine a reasonably contemporary range of studies. Second, we want to provide studies ample time to make it through the peer review and publication process. This date range will be sufficient for both of these goals. However, if more studies are needed in order to obtain sufficient power for our analysis, we will begin rolling the lower-bound date limiter back by one year until a sufficient number of dissertations (and their published versions) is obtained. Two

further delimiters will be used. First, we will use Ph.D. OR Ed.D under the search delimiter category “degree,” and “education” will be used as a subject delimiter.

Once we have obtained the relevant studies for our analysis we will fit a multi-level logistic regression model to examine the odds of an outcome reported in a doctoral dissertation to appear in a peer-reviewed, published article. The first level of the model will be the within-study model, with the dependent variable being whether not a given outcome from the dissertation is reported in the published article. Predictors at this level will include characteristics of the outcome such as the statistical significance level of the treatment comparison, and whether the outcome indicates a “positive” or “negative” result in light of the dissertation’s research goals. The published study report will be the second level of analysis, and we will include relevant study-level covariates such as time elapsed since dissertation completion, sample size, and characteristics of the publication outlet (peer-reviewed or not, etc.).

One important consideration for the problem of outcome reporting bias is the extent to which the problem is hidden. The most egregious cases of outcome reporting bias occur when the study authors fail to mention that an outcome was measured. Less extreme, authors may indicate that an outcome was measured but not give readers any way to determine how large the effect might have been (e.g., a simple statement that results were not statistically significant). We will describe the extent to which each of these patterns occur (and potentially other patterns as well).

An additional consideration for missing outcomes is the extent to which these outcomes are correlated with reported outcomes. For example, assume a study author administers two different self-esteem inventories but reports only the one that revealed a statistically significant relationship. This is not likely to create a large degree of bias, because the two self-esteem

inventories are likely to be highly correlated. Alternatively, assume that a researcher measures self-esteem and academic achievement but only reports results for self-esteem because the results for academic achievement did not reach statistical significance. Here, the two outcomes are not highly correlated, and as such the bias associated with data censoring is likely to be more problematic. To address this issue, we will collect as much descriptive information as possible about the outcomes, including their intercorrelations when available, and will use a Monte Carlo simulation strategy to estimate the potential impact of the data censoring process on conclusions arising from a meta-analysis of a set of studies in which some outcomes not reaching statistical significance are censored. For example, we might simulate study findings for a meta-analysis where each study contributes two outcomes. One of these outcomes will be subject to a censoring mechanism based on the statistical significance of the two-group comparison. Factors to vary in the simulation could then be the correlation between the two outcomes, the fraction of missing outcomes, and the number of studies in the meta-analysis.

### **Conclusions:**

The results of this study will inform both practitioners and researchers by clarifying a currently unknown problem within the field of education. First, this study will provide a foundational understanding for both authors and editors about the magnitude of publication bias in the field of education research. If substantial censorship is found, editorial interventions can then be recommended, crafted, and empirically justified. Researchers will be able to use the information generated from this study to improve reporting standards of primary research, leading to more accurate comprehensive reviews of this literature. Policy-makers will benefit from this study as well. As the impact of publication bias is uncovered in education and as research synthesis continues to be a significant vehicle for educational transformation, policy-

makers will be provided with a more informative culmination of evidence about various educational interventions.

## Appendices

### Appendix A. References

- Chan, A., & Altman, D. G. (2005). Identifying outcome reporting bias in randomised trials on PubMed: Review of publications and survey of authors. *BMJ (Clinical Research Ed.)*, *330*(7494), 753.
- Dickersin, K. (2005). Publication bias: Recognizing the problem, understanding 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. 11-33). Chichester, UK: Wiley.
- Sutton, A. J., & Pigott, T. D. (2005). Bias in meta-analysis induced by incompletely reported studies. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 223-239). Chichester, UK: Wiley.
- 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. *The New England Journal of Medicine*, *358*(3), 252-260.