**Modeling Dependent Effect Sizes in Meta-analysis: Comparing Two Approaches**

Frederick L. Oswald, Chen Zuo, & Evan S. Mulfinger

Rice University

Meta-analysis has become a very common way of summarizing scientific findings in a given organizational research domain; practically every issue of a top-tier organizational journal contains a meta-analysis these days. Across studies, the measures, settings, and samples can differ widely, and meta-analyses attempts to model the heterogeneity in resultant effect sizes with mathematical precision. For example, the Schmidt and Hunter method of meta-analysis applies psychometric corrections to the observed distribution of effect sizes, correcting its mean and variance for measurement error variance, range restriction, and sampling error variance (Schmidt & Hunter, 2014).

In pursuit of even greater mathematical precision, meta-analytic models have accommodated studies reporting *dependent effect sizes*. For example, a study might report correlations between a measure of employee conscientiousness and their task-related job performance at two points in time. These two conscientiousness-performance correlations are dependent because they come from the same sample and they both will be used in the meta-analysis. Also, these two correlations both share the same predictor measure of conscientiousness—however, correlations from the same sample are dependent even when *no* measures are shared (Steiger, 1980).

Meta-analyses can take three different approaches in dealing with this issue. The first approach is to *avoid dependent effects*: Drop one of the correlations (and report that this was done). This would generally be recommended against, although it could be a reasonable decision if this were the only study with dependent effects in a large analysis. The second approach is a common one, which is to *ignore dependent effects*:Average the effects together, treating the averaged effect as if it were independent, like the rest of the effects in the meta-analysis. Averaging is problematic because it ignores known heterogeneity in the effects (in the current example, there is heterogeneity due to time of performance measurement). This results in underestimating the “true” variance across effects in the meta-analysis (i.e., random-effects variance that is not due to psychometric artifacts).

The third approach is the only one that explicitly respects the heterogeneity of all effect sizes, which is to *model dependent effects*: Use a meta-analysis model that estimates heterogeneity *within* sets of dependent effects, which in turn contributes to the estimate of heterogeneity *across* all effects the meta-analysis. Current meta-analysis practice favors random-effects models, which explicitly estimate the “true” variance of the effects across studies (see Borenstein, Hedges, Higgins, & Rothstein, 2009; Cheung, 2008; Hunter & Schmidt, 2000).

Assume that the dependency between multiple effect sizes from a given sample can be reflected in a single number between 0 (completely independent) and 1 (completely dependent). Keeping this assumption in mind, there are two general approaches that characterize how dependent effects in a random-effects meta-analysis are estimated. The first approach might be called *simple*, where the level of effect-size dependency is given a single number—say, .50—that applies to all sets of dependent effect sizes in the meta-analysis. Here, we use just a single parameter estimate that reflects the amount of dependency; this value may be an educated guess or come from the data themselves. The second approach might be called *complex*, where the amount of dependency within each set of dependent effect sizes in the meta-analysis gets estimated (even when there are only two effects coming from the same sample).

The description of these two approaches usefully illustrate the types of methods available for estimating effect-size dependency—even though the approaches oversimplify the methods themselves. A simulation from many years ago (Marín-Martínez & Sánchez-Meca, 1999) compared two specific methods for estimating dependent effects, one relatively more complex (Hedges and Olkin, which estimates the average effect size) than the other (Rosenthal and Rubin, which estimates a composite effect size). This simulation was informative for meta-analysis, but it only focused on averaging 3 dependent effects (*d*-values), only focused on the resultant mean value, and did not include multiple samples.

We will conduct a study comparing dependent effect sizes that expands on this previous simulation by (a) including a much larger range of effect sizes, (b) including a focus on the random-effects variance estimate (tau-squared), in addition to the mean estimate, (c) using both real-world and simulation data, and (d) using more modern methods, namely the robust method of Hedges, Tipton, and Johnson (2010) as the simple approach using the robumeta package in R (Fisher & Tipton, 2015) and SPSS (Tanner-Smith & Tipton, 2014), and Cheung’s structural equation modeling method as the more complex approach (Cheung, 2014b) using the metaSEM package in R (Cheung, 2014a). Results from the real-world data will be compared across these two meta-analysis methods. The simulation results will make a similar comparison, but it also will table the accuracy of meta-analysis results vs. the mean and variance of the population effect sizes that generated the simulation data. This accuracy is indexed by the mean squared error (MSE, which is equal to the squared bias plus the observed variance of the estimate).

This study will provide researchers with practical advice on the best methods for dealing with dependent effect sizes to yield more accurate results in meta-analysis. The study will also conclude with advice on how researchers and journals might better report information on dependent effects in future studies, in anticipation of meta-analyses that involve those effects.

References

Cheung, M. W. -L. (2008). A model for integrating fixed-, random-, and mixed-effects meta-analysis into structural equation modeling. *Psychological Methods, 13,* 182-202.

Cheung, M.W.-L. (2014a). MetaSEM: An R package for meta-analysis using structural equation modeling. *Frontiers in Psycholo*gy, *5*, (1521, 1-7).

Cheung, M. W. -L. (2014b). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods, 19,* 221-229.

Fisher, Z. & Tipton, E. (2015). robumeta: An R-package for robust variance estimation in meta-analysis. <https://arxiv.org/abs/1503.02220>

Hedges, L. V., Tipton, E., Johnson, M. C. (2010) Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1,* 39–65 [Erratum in *1*, 164-165].

Hunter, J. E., & Schmidt, F. L. (2000). Fixed-effects vs. random-effects meta-analysis models: Implications for cumulative knowledge. *International Journal of Selection and Assessment, 8,* 275-292.

Marín-Martínez, F. & Sánchez-Meca, J. (1999). Averaging dependent effect sizes in meta-analysis: A cautionary note about procedures. *The Spanish Journal of Psychology, 2,* 32-38.

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

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87,* 245-251.

Tanner-Smith, E. & Tipton, E. (2014). Robust variance estimation with dependent effect sizes: Practical considerations including a software tutorial in Stata and SPSS. *Research Synthesis Methods, 5,* 13-30.