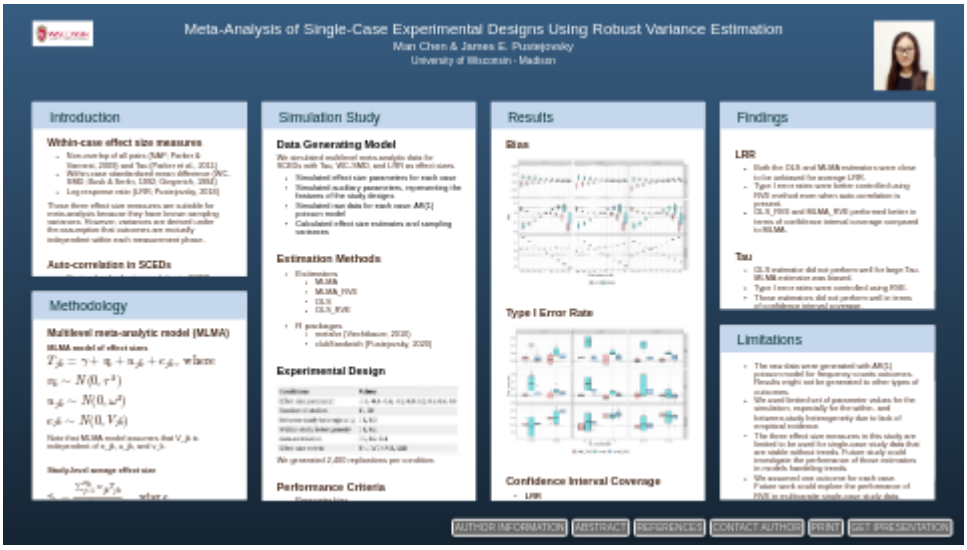
Meta-Analysis of Single-Case Experimental Designs Using Robust Variance Estimation



Man Chen & James E. Pustejovsky

University of Wisconsin - Madison

PRESENTED AT:

INTRODUCTION

Within-case effect size measures

- Non-overlap of all pairs (NAP; Parker & Vannest, 2009) and Tau (Parker et al., 2011)
- Within-case standardized mean difference (WC-SMD; Busk & Serlin, 1992; Gingerich, 1984)
- Log response ratio (LRR; Pustejovsky, 2018)

Those three effect size measures are suitable for meta-analysis because they have known sampling variances. However, variances are derived under the assumption that outcomes are mutually independent within each measurement phase.

Auto-correlation in SCEDs

- Varying levels of auto-correlation in SCEDs (Shadish & Sullivan, 2011; Solomon, 2014; Barnard-Brak, 2021)
- Consequences
  - Positive auto-correlation leads to under-estimated sampling variances of the effect size estimates.
  - Results in inaccurate estimates of the overall average effect size and standard error.
  - Leads to inappropriate statistical inference about the intervention effects.

Purpose of this study

- We use simulation to evaluate the performance of multilevel meta-analysis, with and without use of robust variance estimation (RVE), in the presence of auto-correlation when using NAP/Tau, WC-SMD, and LRR for meta-analyzing SCEDs.
- We compare MLMA model to OLS, a simpler alternative.

METHODOLOGY

Multilevel meta-analytic model (MLMA)

MLMA model of effect sizes

$$T_{jk} = \gamma + v_k + u_{jk} + e_{jk}, \text{ where}$$
$$v_k \sim N(0, \tau^2)$$
$$u_{jk} \sim N(0, \omega^2)$$
$$e_{jk} \sim N(0, V_{jk})$$

Note that MLMA model assumes that V_jk is independent of e_jk, u_jk, and v_k.

Study-level aveage effect size

$$\hat{\gamma}_k = \frac{\sum_{j=1}^{n_k} w_{jk} T_{jk}}{\sum_{j=1}^{n_k} w_{jk}}, \quad \text{where} \quad w_{jk} = \frac{1}{\hat{\omega}^2 + V_{jk}}$$

Overall average effect size

$$\hat{\gamma} = \frac{\sum_{k=1}^{K} h_k \hat{\gamma}_k}{H}, \quad \text{where}$$
$$H = \sum_{k=1}^{K} h_k,$$
$$h_k = 1 \Bigg/ \left(\hat{\tau}^2 + \frac{1}{\sum_{j=1}^{n_k} w_{jk}} \right)$$

MLMA_RVE

- Pustejovsky (2018) proposed the use of RVE in conjunction with MLMA of SCEDs.
- To address the possibility that the samping variances of effect sizes might be estimated inaccurately due to auto-correlation.
- To produce standard errors that take into account the possibility of within-case errors and the dependence structure arising from having multiple effect sizes nested within studies.

$$V^R = \frac{\sum_{k=1}^{K} h_k^2 (\hat{\gamma}_k - \hat{\gamma})^2}{H^2}, \quad V^{CR} = \frac{1}{H^2} \sum_{k=1}^{K} \frac{h_k^2 (\hat{\gamma}_k - \hat{\gamma})^2}{1 - \frac{h_k}{H}}$$

Ordinary least squares (OLS)

- Easy to compute
- Does not require that V_jk is independent of e_jk, u_jk, and v_k.

$$\bar{\gamma} = \frac{\sum_{k=1}^{K} \sum_{j=1}^{n_k} T_{jk}}{N},$$
$$V_{ols}^R = \frac{\sum_{k=1}^{K} n_k^2 (\bar{\gamma}_k - \bar{\gamma})^2}{N^2},$$
$$V_{ols}^{CR} = \frac{1}{N^2} \sum_{k=1}^{K} \frac{n_k^2 (\bar{\gamma}_k - \bar{\gamma})^2}{1 - \frac{n_k}{N}}$$

SIMULATION STUDY

Data Generating Model

We simulated multilevel meta-analytic data for SCEDs with Tau, WC-SMD, and LRR as effect sizes.

- Simulated effect size parameters for each case
- Simulated auxiliary parameters, representing the features of the study designs
- Simulated raw data for each case: AR(1) poisson model
- Calculated effect size estimates and sampling variances

Estimation Methods

- Estimators
  - MLMA
  - MLMA_RVE
  - OLS
  - OLS_RVE
- R packages
  - metafor (Viechtbauer, 2010)
  - clubSandwich (Pustejovsky, 2020)

Experimental Design

Conditions	Values
Effect size parameter	-0.8, -0.6, -0.4, -0.2, 0.0, 0.2, 0.4, 0.6, 0.8
Number of studies	10, 30
Between-study heterogeneity	0.1, 0.3
Within-study heterogeneity	0.1, 0.2
Auto-correlation	0.0, 0.2, 0.4
Effect size metric	Tau, WC-SMD, LRR

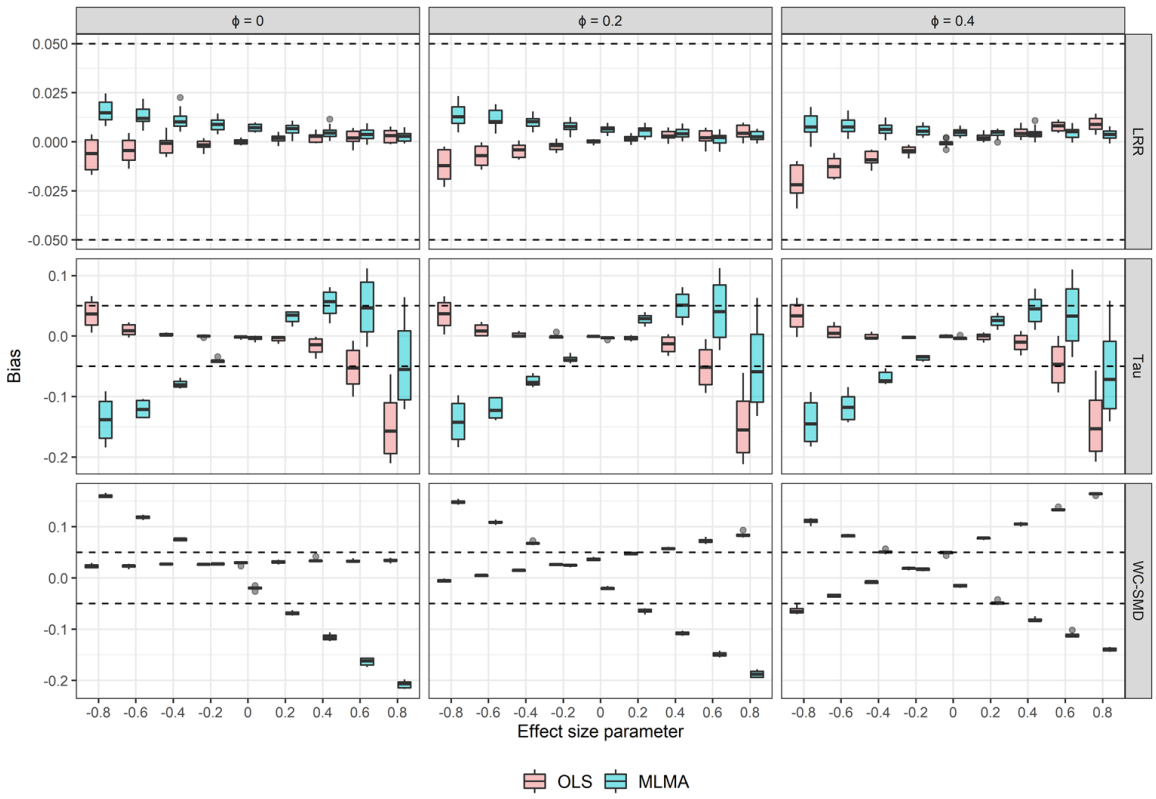
We generated 2,400 replications per condition.

Performance Criteria

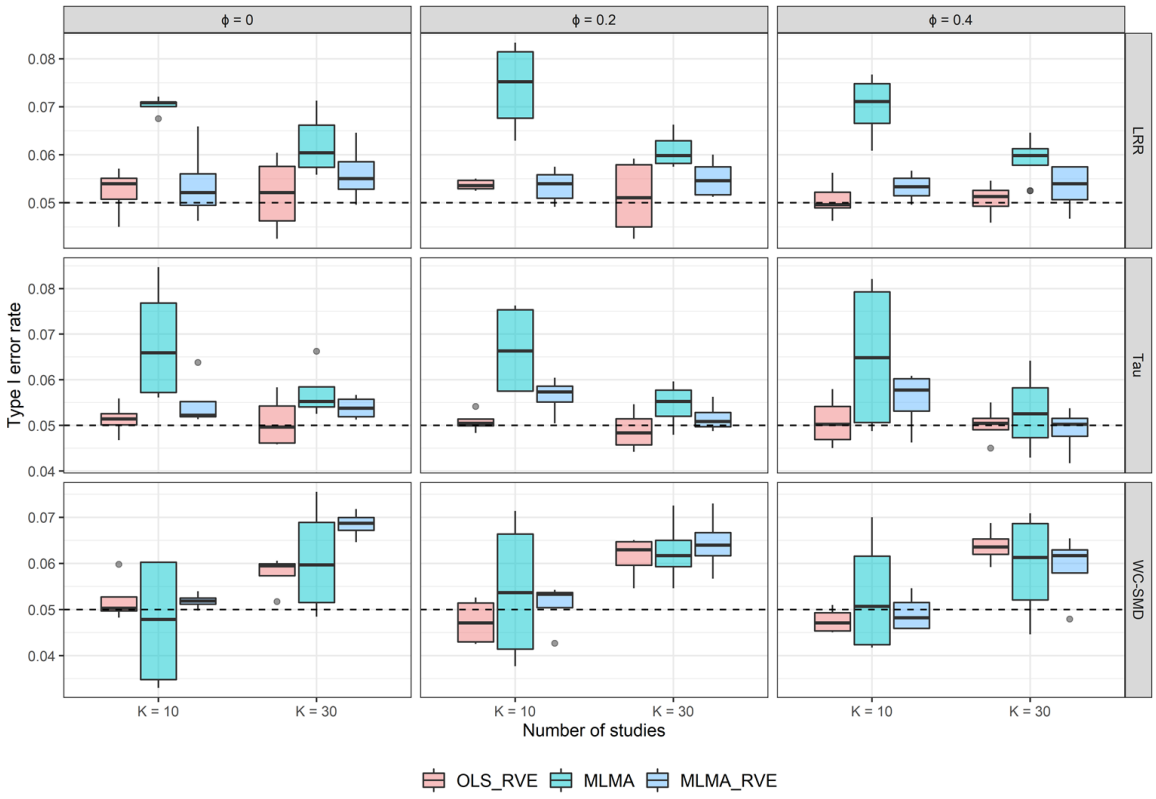
- Parameter bias
- Type I error rate
- Confidence interval converage

RESULTS

Bias

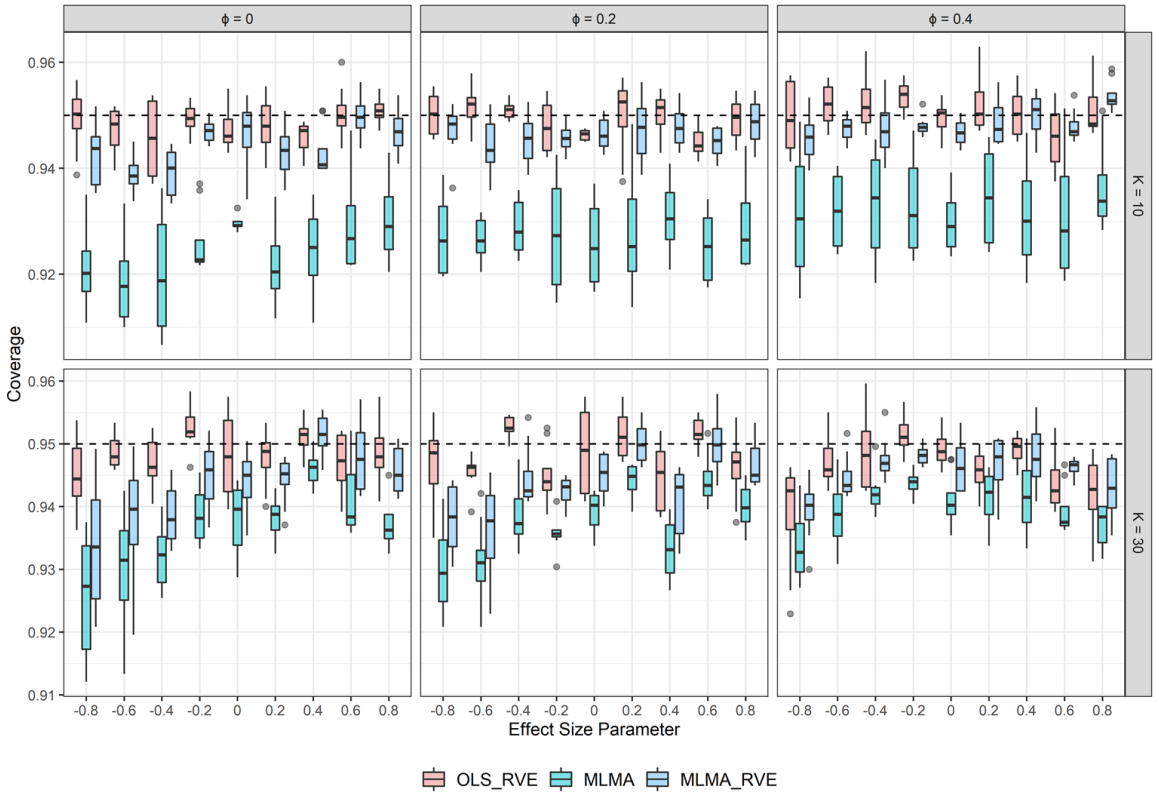


Type I Error Rate

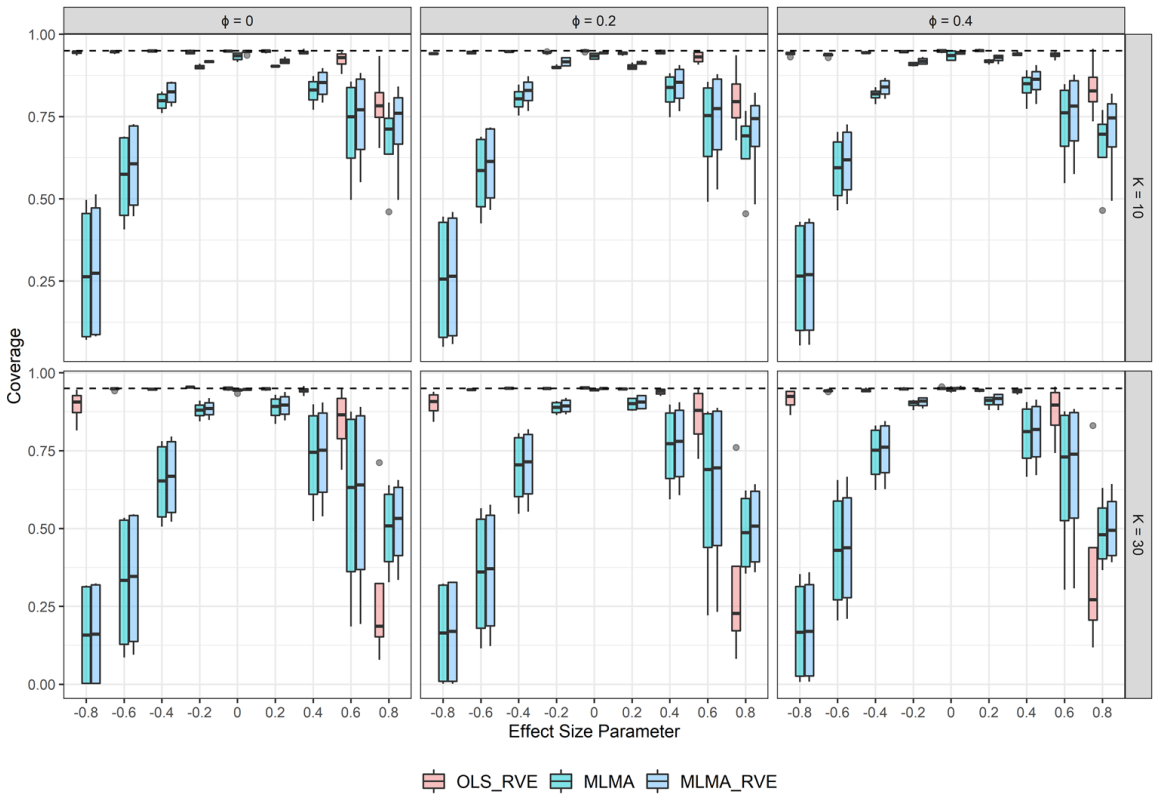


Confidence Interval Coverage

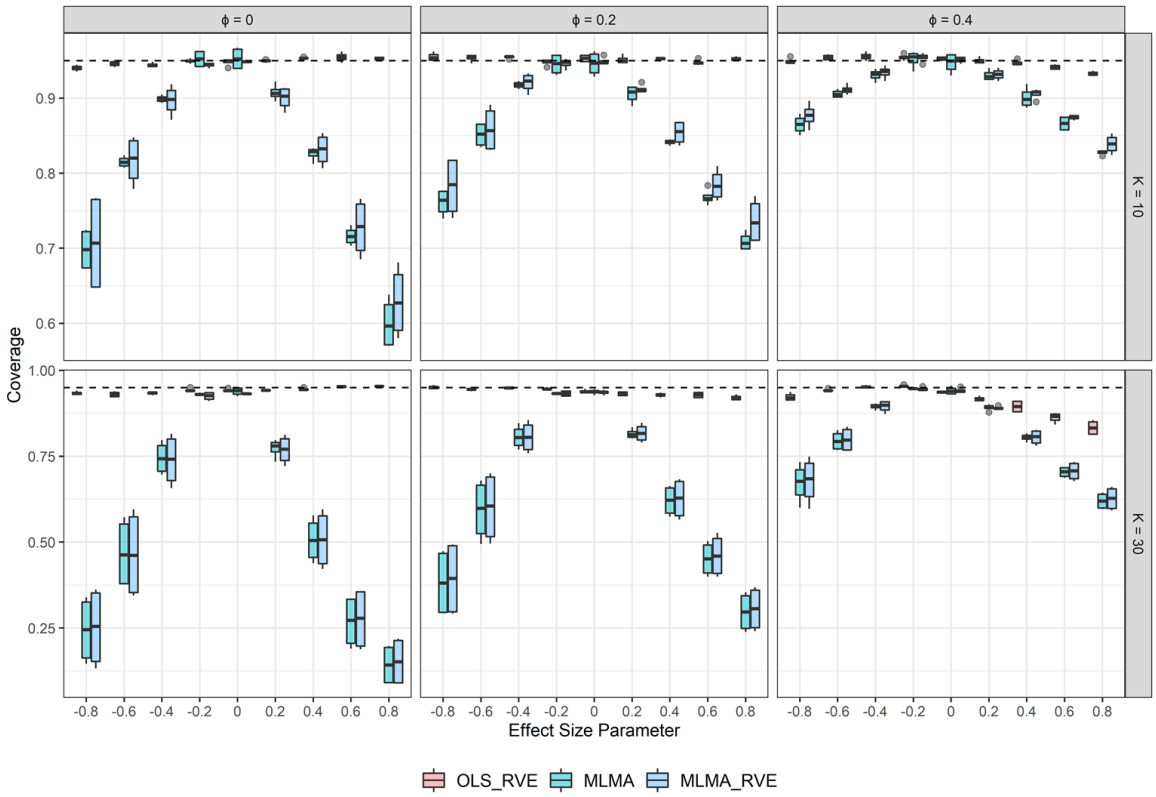
- **LRR**



- **Tau**


ϕ = 0
ϕ = 0.2
ϕ = 0.4
K = 10
K = 30
Coverage
Effect Size Parameter
OLS_RVE
MLMA
MLMA_RVE

- **WC-SMD**


ϕ = 0
ϕ = 0.2
ϕ = 0.4
K = 10
K = 30
Coverage
Effect Size Parameter
OLS_RVE
MLMA
MLMA_RVE

FINDINGS

LRR

- Both the OLS and MLMA estimators were close to be unbiased for average LRR.
- Type I error rates were better controlled using RVE method even when auto-correlation is present.
- OLS_RVE and MLMA_RVE performed better in terms of confidence interval coverage compared to MLMA.

Tau

- OLS estimator did not perform well for large Tau. MLMA estimator was biased.
- Type I error rates were controlled using RVE.
- Those estimators did not perform well in terms of confidence interval coverage.

WC-SMD

- OLS estimator was biased when there is auto-correlation. MLMA estimator was systematically biased.
- Type I error rates were not controlled when the number of studies was large.
- OLS_RVE performed better in terms of confidence interval coverage compared to MLMA and MLMA_RVE but its performance was impacted by auto-correlation.

LIMITATIONS

- The raw data were generated with AR(1) poisson model for frequency counts outcomes. Results might not be generated to other types of outcomes.
- We used limited set of parameter values for the simulation, especially for the within- and between-study heterogeneity due to lack of empirical evidence.
- The three effect size measures in this study are limited to be used for single-case study data that are stable without trends. Future study could investigate the performance of those estimators in models handeling trends.
- We assumed one outcome for each case. Future work could explore the performance of RVE in multivaraite single-case study data.

AUTHOR INFORMATION

Man Chen (she/her/hers)

Doctoral student, Quantitative Methods

Department of Educational Psychology

University of Wisconsin-Madison

Email: mchen277@wisc.edu

Dr. James E. Pustejovsky

Associate Professor

Department of Educational Psychology

University of Wisconsin-Madison

ABSTRACT

Single-case experimental designs (SCEDs) involve measuring an outcome repeatedly over time across two or more distinct phases, such as a baseline phase and an intervention phase. With such data, the effect of intervention is assessed by comparing the data series during the baseline phase with the series during the treatment phase(s) within each case, so that each participant serves as their own control (Horner & Odom, 2014). In bodies of literature where there are many SCEDs examining a common class of intervention, meta-analysis methods may be used to synthesize results across studies and draw conclusions about the overall effects of intervention, about variability in effects, and about factors associated with larger or smaller effects (Pustejovsky and Ferron, 2017). Because of the structure of SCEDs, in which multiple cases are nested within studies, syntheses of SCED should make use of multilevel meta-analytic models (Van den Noortgate & Onghena, 2008). One approach to quantifying the direction and magnitude of intervention effect for SCEDs is through within-case effect size measures. Among the within-case effect size measures, the non-overlap of all pairs (NAP: Parker & Vannest, 2009), the within-case standardized mean difference (WC-SMD: Busk & Serlin, 2015: Gingerich, 1984), and the log response ratio (LRR: Pustejovsky, 2018) are suitable for meta-analysis because they have known sampling variances, allowing researchers to make statistical inference about the intervention effects. Available formulas for the sampling variances of NAP, WC-SMD, and LRR are all based on the assumption that outcome measurements are independent within each measurement phase. This assumption would be violated if outcomes are instead auto-correlated—a plausible concern given that the outcomes represent repeated measures on a single unit. In the presence of positive autocorrelation, the sampling variances will be estimated downward, which could lead to inflated type I error rate and inaccurate statistical inferences (Beretvas & Chung, 2008: Petit-Bois, et al., 2016). However, the implications of auto-correlation for meta-analysis of these effect sizes has yet to be examined. Robust variance estimation (RVE, Hedges, Tipton, & Johnson, 2010: Tipton, 2015: Tipton & Pustejovsky, 2015) has been widely used in the meta-analysis of group experimental designs to estimate unbiased standard errors for overall average treatment effect estimates and meta-regression coefficients. RVE has been proposed as a solution for meta-analysis of SCEDs when sampling variances of effect sizes might be estimated inaccurately due to autocorrelation (Pustejovsky, 2018), but its performance has yet to be assessed. In this study, we use simulation to evaluate the performance of multi-level meta-analysis, with and without use of RVE, in the presence of autocorrelation when using NAP, WC-SMD, and LRR for meta-analyzing SCEDs. We examine the bias of overall average effect size estimates, standard errors, and variance component estimates, as well as the calibration of confidence intervals for overall average effect sizes.

REFERENCES

Barnard-Brak, L., Watkins, L., & Richman, D. M. Autocorrelation and estimates of treatment effect size for single-case experimental design data. Behavioral Interventions.

Beretvas, S. N., & Chung, H. (2008). An evaluation of modified r^2-change effect size indices for single-subject experimental designs. *Evidence-Based Communication Assessment and Intervention, 2* (3), 120–128.

Busk, P. L., & Serlin, R. C. (1992). Meta-analysis for single-case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research design and analysis: New directions for psychology and education* (pp. 187–212). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Collins, T. A., Drevon, D. D., Brown, A. M., Villarreal, J. N., Newman, C. L., & Endres, B. (2020). Say something nice: A meta-analytic review of peer reporting interventions. *Journal of School Psychology, 83*, 89–103.

Gingerich, W. J. (1984). Meta-analysis of applied time-series data. *Journal of Applied Behavioral Science, 20* (1), 71–79. https://doi.org/10.1177/002188638402000113

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.

Horner, R. H., & Odom, S. L. (2014). Constructing single-case research designs: Logic and options. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case intervention research: Methodological and statistical advances* (pp. 27–51). Washington, DC: American Psychological Association.

Maggin, D. M., Pustejovsky, J. E., & Johnson, A. H. (2017). A meta-analysis of school-based group contingency interventions for students with challenging behavior: An update. *Remedial and Special Education, 38* (6), 353–370.

Moeyaert, M., Ugille, M., Ferron, J. M., Beretvas, S. N., & Van den Noortgate, W. (2013). The three-level synthesis of standardized single-subject experimental data: A monte carlo simulation study. *Multivariate Behavioral Research, 48* (5), 719–748. https://doi.org/10.1080/00273171.2013.816621

Parker, R. I., & Vannest, K. J. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy, 40* (4), 357–367.

Petit-Bois, M., Baek, E. K., Van den Noortgate, W., Beretvas, S. N., & Ferron, J. M. (2016). The consequences of modeling autocorrelation when synthesizing single-case studies using a three-level model. *Behavior Research Methods, 48* (2), 803–812.

Pustejovsky, J. E., & Ferron, J. M. (2017). Research synthesis and meta-analysis of single-case designs. In J. M. Kauffman, D. P. Hallahan, & P. C. Pullen (Eds.), *Handbook of special education* (pp. 168–186). Routledge New York, NY.

Pustejovsky, J. E. (2018). Using response ratios for meta-analyzing single-case designs with behavioral outcomes. *Journal of School Psychology, 68*, 99–112.

Pustejovsky, J. E. (2020). *clubSandwich: Cluster-robust (sandwich) variance estimators with small-sample corrections*. Retrieved from https://github.com/jepusto/clubSandwich

R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Shadish, W. R., & Sullivan, K. J. (2011). Characteristics of single-case designs used to assess intervention effects in 2008. *Behavior Research Methods, 43* (4), 971–980.

Solomon, B. G. (2014). Violations of assumptions in school-based single-case data: Implications for the selection and interpretation of effect sizes. *Behavior Modification, 38* (4), 477–496.

Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods, 20* (3), 375.

Tipton, E., & Pustejovsky, J. E. (2015). Small-sample adjustments for tests of moderators and model fit using robust variance estimation in meta-regression. *Journal of Educational and Behavioral Statistics, 40* (6), 604–634.

Van den Noortgate, W., & Onghena, P. (2008). A multilevel meta-analysis of single-subject experimental design studies. *Evidence-Based Communication Assessment and Intervention, 2* (3), 142–151. https://doi.org/10.1080/17489530802505362

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36* (3), 1–48. Retrieved from https://www.jstatsoft.org/v36/i03/