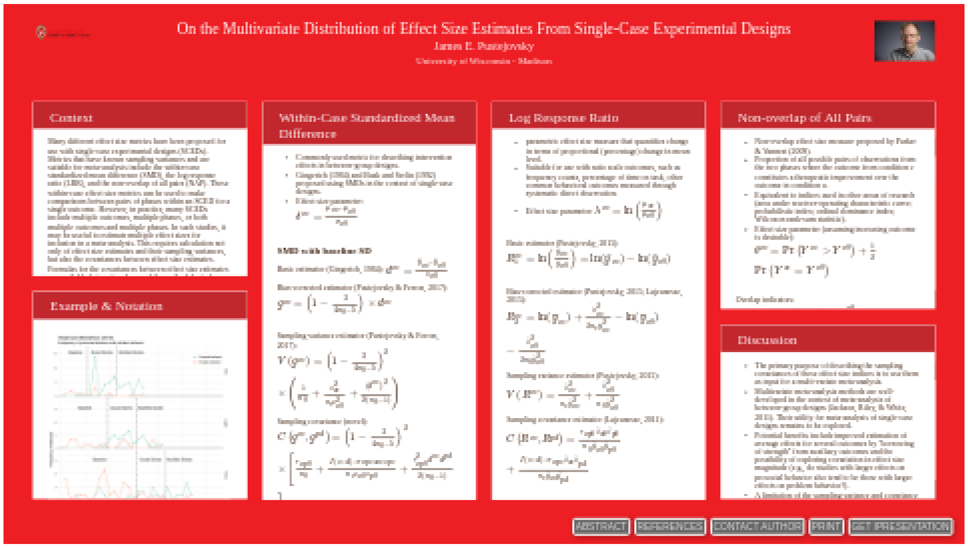
On the Multivariate Distribution of Effect Size Estimates From Single-Case Experimental Designs



James E. Pustejovsky

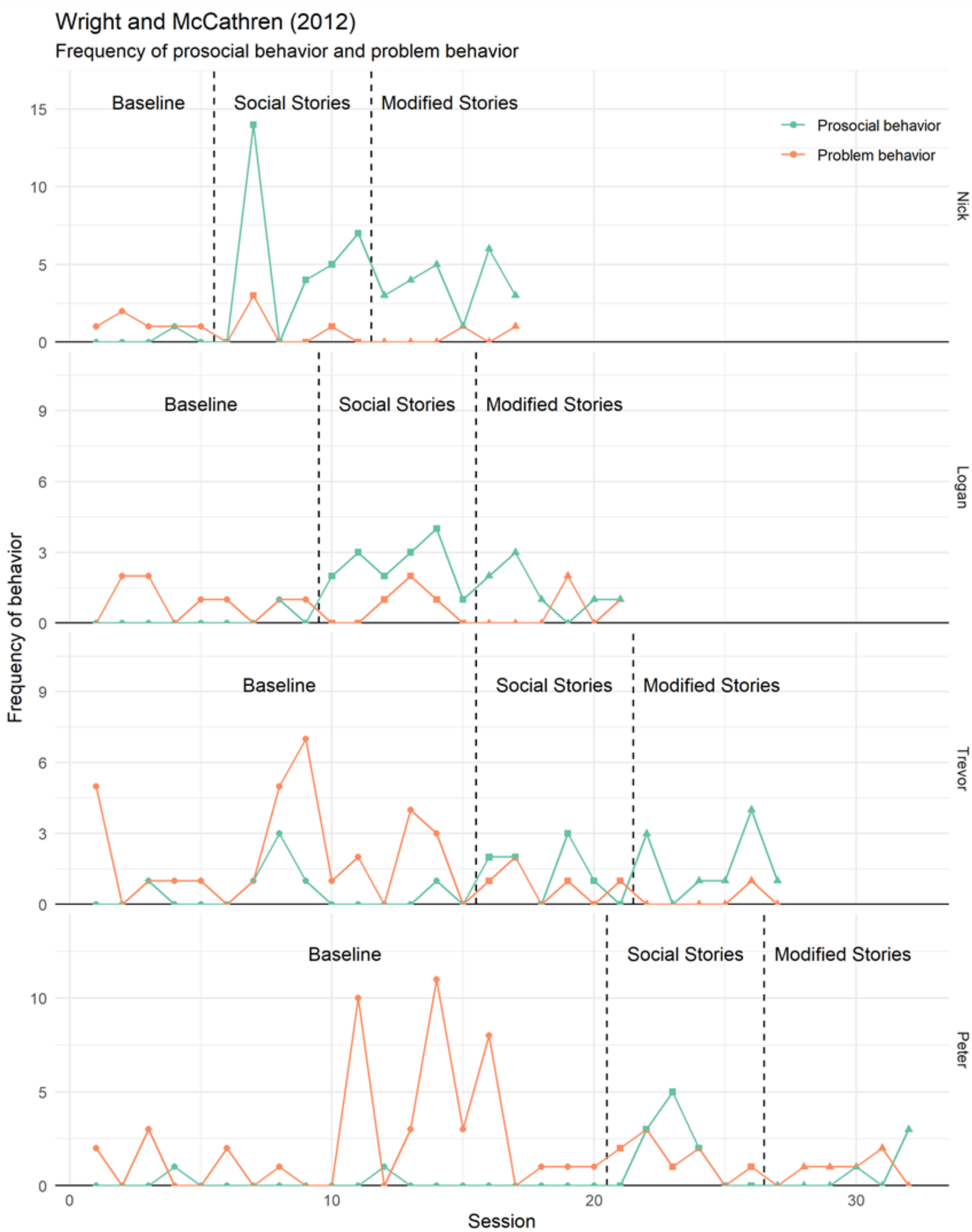
University of Wisconsin - Madison

PRESENTED AT:

CONTEXT

Many different effect size metrics have been proposed for use with single-case experimental designs (SCEDs). Metrics that have known sampling variances and are suitable for meta-analysis include the within-case standardized mean difference (SMD), the log response ratio (LRR), and the non-overlap of all pairs (NAP). These within-case effect size metrics can be used to make comparisons between pairs of phases within an SCED for a single outcome. However, in practice, many SCEDs include multiple outcomes, multiple phases, or both multiple outcomes and multiple phases. In such studies, it may be useful to estimate multiple effect sizes for inclusion in a meta-analysis. This requires calculation not only of effect size estimates and their sampling variances, but also the covariances between effect size estimates. Formulas for the covariances between effect size estimates are available but scattered around the methodological literature. This paper reviews and consolidates available formulas and demonstrates their relevance in the context of meta-analysis of SCEDs. I describe methods for estimating multiple effect sizes, along with corresponding sampling variances and covariances, for the within-case SMD, LRR, and NAP indices. An empirical example is included to illustrate the calculations.

EXAMPLE & NOTATION



Using an across-participant multiple baseline design, Wright and McCathren (2012) examined the effects of a Social Stories intervention on the prosocial behavior and problem behavior of three children with autism. Their study design included three phases: a baseline phase, an initial Social Stories intervention phase, and a modified Social Stories intervention phase. I demonstrate how to calculate effect size estimates, with corresponding sampling variances and covariances, for comparisons of each of the Social Stories phases versus the baseline phase, for both prosocial and problem behavior outcomes.

Notation

Consider a data series consisting of O outcome variables, measured repeatedly under a baseline condition/phase and under C intervention conditions/phases.

n_c number of observations in condition c
y_i^{oc} response i on outcome o in condition c
μ_{oc} mean level of outcome o in condition c
σ_{oc} SD of outcome o in condition c
σ_{opc} covariance of outcomes o and p in condition c

Sample statistics

$\bar{y}_{oc}$ sample mean of outcome o in condition c
s_{oc} sample SD of outcome o in condition c
r_{opc} correlation of outcomes o and p in condition c

$$s_{oP}^2 = \frac{1}{v}\sum_{c=0}^{C}(n_c - 1)s_{oc}^2 \text{ pooled sample variance}$$

$$v = \sum_{c=0}^{C}(n_c - 1) \text{ pooled degrees of freedom}$$

WITHIN-CASE STANDARDIZED MEAN DIFFERENCE

- Commonly used metric for describing intervention effects in between-group designs.
- Gingerich (1984) and Busk and Serlin (1992) proposed using SMDs in the context of single-case designs.
- Effect size parameter:
$\delta^{oc} = \frac{\mu_{oc} - \mu_{o0}}{\sigma_{o0}}$

SMD with baseline SD

Basic estimator (Gingerich, 1984): $d^{oc} = \frac{\bar{y}_{oc} - \bar{y}_{o0}}{s_{o0}}$

Bias-corrected estimator (Pustejovsky & Ferron, 2017): $g^{oc} = \left(1 - \frac{3}{4n_0 - 5}\right) \times d^{oc}$

Sampling variance estimator (Pustejovsky & Ferron, 2017): $V\left(g^{oc}\right) = \left(1 - \frac{3}{4n_0 - 5}\right)^2 \times \left(\frac{1}{n_0} + \frac{s_{oc}^2}{n_c s_{o0}^2} + \frac{(d^{oc})^2}{2(n_0 - 1)}\right)$

Sampling covariance (novel): $C\left(g^{oc}, g^{pd}\right) = \left(1 - \frac{3}{4n_0 - 5}\right)^2 \times \left[\frac{r_{op0}}{n_0} + \frac{I(c=d) \times r_{opc} s_{oc} s_{pc}}{n_c s_{o0} s_{p0}} + \frac{r_{op0}^2 d^{oc} d^{pd}}{2(n_0 - 1)}\right]$

SMD with pooled SD

Basic estimator (Busk & Serlin, 1992): $\tilde{d}^{oc} = \frac{\bar{y}_{oc} - \bar{y}_{o0}}{\sqrt{s_{oP}^2}}$

Bias-corrected estimator (Hedges, 1981): $\tilde{g}^{oc} = \left(1 - \frac{3}{4v - 1}\right) \times \tilde{d}^{oc}$

Sampling variance (Hedges, 1981):

$$V\left(\tilde{g}^{oc}\right) = \left(1 - \frac{3}{4v - 1}\right)^2 \times \left(\frac{1}{n_0} + \frac{1}{n_c} + \frac{\left(\tilde{d}^{oc}\right)^2}{2v}\right)$$

Sampling covariance estimator (Gleser & Olkin, 2009):

$$C\left(\tilde{g}^{oc}, \tilde{g}^{pd}\right) = \left(1 - \frac{3}{4v - 1}\right)^2 \times \left[r_{op}\left(\frac{1}{n_0} + \frac{I(c=d)}{n_c}\right) + \frac{r_{op}^2 d^{oc} d^{pd}}{2v}\right]$$

Wright & McCathren (2012)

Estimated within-case SMDs for Wright & McCathren (2012)

Phase	Outcome	g	SE(g)	V1	V2	V3	V4
Nick							
Modified Stories	Problem	-1.55	0.76	0.57	-0.46	0.31	-0.14
	Prosocial	6.20	2.56	-0.24	6.57	-0.08	6.78
Social Stories	Problem	-0.95	1.01	0.41	-0.03	1.02	2.78
	Prosocial	8.59	4.88	-0.04	0.54	0.56	23.85
Logan							
Modified Stories	Problem	-0.45	0.51	0.26	-0.31	0.10	0.00
	Prosocial	3.31	1.44	-0.42	2.08	0.00	1.43
Social Stories	Problem	-0.26	0.49	0.39	0.01	0.24	0.21
	Prosocial	6.47	2.01	0.00	0.49	0.21	4.06
Trevor							
Modified Stories	Problem	-0.81	0.30	0.09	0.06	0.07	0.02
	Prosocial	1.36	0.78	0.25	0.61	0.02	0.11
Social Stories	Problem	-0.53	0.29	0.86	0.10	0.09	0.06
	Prosocial	0.98	0.64	0.12	0.21	0.33	0.41
Peter							
Modified Stories	Problem	-0.42	0.24	0.06	-0.08	0.05	-0.01
	Prosocial	1.77	1.58	-0.21	2.51	-0.01	0.27
Social Stories	Problem	-0.23	0.25	0.81	-0.03	0.06	0.08
	Prosocial	4.89	2.76	-0.02	0.06	0.11	7.59

Note:
Variance and covariance estimates are in blue. Estimated correlations between effect size estimates are in yellow.

LOG RESPONSE RATIO

- parametric effect size measure that quantifies change in terms of proportional (percentage) change in mean level.
- Suitable for use with ratio scale outcomes, such as frequency counts, percentage of time on task, other common behavioral outcomes measured through systematic direct observation.
- Effect size parameter: $\lambda^{oc} = \ln\left(\frac{\mu_{oc}}{\mu_{o0}}\right)$

Basic estimator (Pustejovsky, 2015): $R_1^{oc} = \ln\left(\frac{\tilde{y}_{oc}}{\tilde{y}_{o0}}\right) = \ln(\tilde{y}_{oc}) - \ln(\tilde{y}_{o0})$

Bias-corrected estimator (Pustejovsky, 2015; Lajeunesse, 2015):

$$R_2^{oc} = \ln(\tilde{y}_{oc}) + \frac{\tilde{s}_{oc}^2}{2n_c\tilde{y}_{oc}^2} - \ln(\tilde{y}_{o0}) - \frac{\tilde{s}_{o0}^2}{2n_0\tilde{y}_{o0}^2}$$

Sampling variance estimator (Pustejovsky, 2015):

$$V\left(R^{oc}\right) = \frac{\tilde{s}_{oc}^2}{n_c\tilde{y}_{oc}^2} + \frac{\tilde{s}_{o0}^2}{n_0\tilde{y}_{o0}^2}$$

Sampling covariance estimator (Lajeunesse, 2011):

$$C\left(R^{oc}, R^{pd}\right) = \frac{r_{op0}\tilde{s}_{o0}\tilde{s}_{p0}}{n_0\tilde{y}_{o0}\tilde{y}_{p0}} + \frac{I(c=d)\times r_{opc}\tilde{s}_{oc}\tilde{s}_{pd}}{n_c\tilde{y}_{oc}\tilde{y}_{pd}}$$

Wright & McCathren (2012)

Estimated LRRs for Wright & McCathren (2012)

Phase	Outcome	R	SE(R)	V1	V2	V3	V4
Nick							
Modified Stories	Problem	-1.09	0.65	0.43	-0.13	0.03	-0.04
	Prosocial	2.43	1.02	-0.20	1.04	-0.04	1.00
Social Stories	Problem	-0.33	0.76	0.06	-0.05	0.58	0.23
	Prosocial	2.81	1.09	-0.06	0.90	0.28	1.18
Logan							
Modified Stories	Problem	-0.38	0.74	0.55	-0.13	0.09	0.02
	Prosocial	2.03	1.05	-0.17	1.10	0.02	1.00
Social Stories	Problem	-0.21	0.58	0.20	0.03	0.34	0.06
	Prosocial	2.63	1.01	0.02	0.94	0.09	1.03
Trevor							
Modified Stories	Problem	-2.06	1.04	1.08	0.34	0.08	0.06
	Prosocial	1.23	0.59	0.56	0.35	0.06	0.21
Social Stories	Problem	-0.88	0.46	0.16	0.23	0.21	0.13
	Prosocial	1.01	0.59	0.10	0.61	0.48	0.35
Peter							
Modified Stories	Problem	-1.00	0.49	0.24	-0.19	0.11	-0.05
	Prosocial	1.94	1.01	-0.39	1.02	-0.05	0.47
Social Stories	Problem	-0.44	0.44	0.50	-0.12	0.19	-0.01
	Prosocial	2.70	0.85	-0.12	0.55	-0.03	0.73

Note:
Variance and covariance estimates are in blue. Estimated correlations between effect size estimates are in yellow.

NON-OVERLAP OF ALL PAIRS

- Non-overlap effect size measure proposed by Parker & Vannest (2009).
- Proportion of all possible pairs of observations from the two phases where the outcome from condition *c* constitutes a therapeutic improvement over the outcome in condition *o*.
- Equivalent to indices used in other areas of research (area under receiver-operating characteristic curve; probabilistic index; ordinal dominance index; Wilcoxon rank-sum statistic).
- Effect size parameter (assuming increasing outcome is desirable):
  $\theta^{oc} = \Pr\left(Y^{oc} > Y^{o0}\right) + \frac{1}{2}\Pr\left(Y^{oc} = Y^{o0}\right)$

Overlap indicators:

$$q_{ij}^{oc} = \begin{cases} 1 & \text{if} \quad y_j^{oc} > y_i^{o0} \\ \frac{1}{2} & \text{if} \quad y_j^{oc} = y_i^{o0} \\ 0 & \text{if} \quad y_j^{oc} < y_i^{o0} \end{cases}$$

Estimator (Parker & Vannest, 2009):

$$\hat{\theta}^{oc} = \frac{1}{n_0 n_c}\sum_{i=1}^{n_0}\sum_{j=1}^{n_c} q_{ij}^{oc}$$

Sampling variance estimator (Sen, 1967; Mee, 1990):

$$V\left(\hat{\theta}^{oc}\right) = \frac{1}{(n_0-1)(n_c-1)}\left[\hat{\theta}^{oc}(1-\hat{\theta}^{oc}) + n_c Q_A^{oc} + n_0 Q_B^{oc} - 2Q_C^{oc}\right]$$

where

$$Q_A^{oc} = \frac{1}{n_0 n_c^2}\sum_{i=1}^{n_0}\left[\sum_{j=1}^{n_c}\left(q_{ij}^{oc} - \hat{\theta}^{oc}\right)\right]^2$$

$$Q_B^{oc} = \frac{1}{n_0^2 n_c}\sum_{j=1}^{n_c}\left[\sum_{i=1}^{n_0}\left(q_{ij}^{oc} - \hat{\theta}^{oc}\right)\right]^2$$

$$Q_C^{oc} = \frac{1}{n_0 n_c}\sum_{i=1}^{n_0}\sum_{j=1}^{n_c}\left(q_{ij}^{oc} - \hat{\theta}^{oc}\right)^2$$

Sampling covariance estimator (Delong, Delong, & Clarke-Pearson, 1988):

$$C\left(\hat{\theta}^{oc}, \hat{\theta}^{pd}\right) = \frac{Q_A^{opcd}}{n_0-1} + \frac{I(c=d)\times Q_B^{opc}}{n_c-1}$$

$$Q_A^{opcd} = \frac{1}{n_0}\sum_{i=1}^{n_0}\left[\frac{1}{n_c}\sum_{j=1}^{n_c}\left(q_{ij}^{oc} - \hat{\theta}^{oc}\right)\right]\left[\frac{1}{n_d}\sum_{j=1}^{n_d}\left(q_{ij}^{pd} - \hat{\theta}^{pd}\right)\right]$$

$$Q_B^{opc} = \frac{1}{n_0^2 n_c}\sum_{j=1}^{n_c}\left[\sum_{i=1}^{n_0}\left(q_{ij}^{oc} - \hat{\theta}^{oc}\right)\right]\left[\sum_{i=1}^{n_0}\left(q_{ij}^{pc} - \hat{\theta}^{pc}\right)\right]$$

Wright & McCathren (2012)

Estimated NAPs for Wright & McCathren (2012)

Condition	Outcome	NAP	SE(NAP)	V1	V2	V3	V4
Nick							
Modified Stories	Problem	0.13	0.09	0.008	-0.001	0.001	-0.000
	Prosocial	0.98	0.02	-0.481	0.001	-0.000	0.001
Social Stories	Problem	0.23	0.17	0.037	-0.018	0.028	0.009
	Prosocial	0.80	0.13	-0.023	0.180	0.420	0.017
Logan							
Modified Stories	Problem	0.35	0.15	0.022	-0.009	0.007	0.000
	Prosocial	0.88	0.10	-0.645	0.009	0.000	0.000
Social Stories	Problem	0.42	0.15	0.323	0.027	0.023	0.001
	Prosocial	0.99	0.01	0.060	0.275	0.281	0.000
Trevor							
Modified Stories	Problem	0.18	0.08	0.007	0.004	0.005	0.001
	Prosocial	0.78	0.11	0.398	0.012	0.002	0.003
Social Stories	Problem	0.37	0.12	0.539	0.163	0.014	0.007
	Prosocial	0.72	0.13	0.120	0.190	0.405	0.018
Peter							
Modified Stories	Problem	0.42	0.11	0.013	-0.004	0.007	-0.001
	Prosocial	0.62	0.11	-0.293	0.013	-0.001	0.000
Social Stories	Problem	0.55	0.12	0.524	-0.060	0.014	0.005
	Prosocial	0.72	0.12	-0.039	0.035	0.326	0.015

Note:
Variance and covariance estimates are in blue.
Estimated correlations between effect size estimates are in yellow.

DISCUSSION

- The primary purpose of describing the sampling covariances of these effect size indices is to use them as input for a multi-variate meta-analysis.
- Multivariate meta-analysis methods are well-developed in the context of meta-analysis of between-group designs (Jackson, Riley, & White, 2011). Their utility for meta-analysis of single-case designs remains to be explored.
- Potential benefits include improved estimation of average effects for several outcomes by "borrowing of strength" from auxiliary outcomes and the possibility of exploring covariation in effect size magnitude (e.g., do studies with larger effects on prosocial behavior also tend to be those with larger effects on problem behavior?).
- A limitation of the sampling variance and covariance formulas is that they are based on the assumption that outcomes within each condition/phase are mutually independent, rather than auto-correlated.
- It would be valuable to extend these methods to allow for an estimated (or fixed, assumed) value of AR(1) auto-correlation.

ABSTRACT

Single-case experimental designs (SCEDs) are a class of study designs that use repeated measurement of outcomes on one or a small number of cases in order to evaluate the effects of an intervention. SCEDs comprise a substantial fraction of the evidence base in certain research areas—particularly on topics such as interventions for low-incidence populations, where it is difficult to obtain samples large enough to use between-subjects research designs. Consequently, there is a need for methods to synthesize findings from SCEDs, such as through the use of meta-analysis methods. A variety of methods have been developed for estimating effect sizes and conducting meta-analysis of data from SCEDs, including methods based on within-case effect size metrics, between-case effect size metrics, or raw data synthesis (Pustejovsky & Ferron, 2017). Approaches based on within-case effect size metrics involves first estimating effect sizes for each case in each study, then synthesizing these estimates using multi-level meta-analysis (Van den Noortgate & Onghena, 2008). This generate estimates of an overall average effect, as well as estimates of within- and between-study heterogeneity. Many different effect size metrics have been proposed for use with single-case designs (Manolov & Moeyaert, 2017: Parker, Vannest, & Davis, 2014). Metrics that have known sampling variances and are suitable for meta-analysis include the within-case standardized mean difference (SMD: Gingerich, 1984), the log response ratio (LRR: Pustejovsky, 2015), and the non-overlap of all pairs (NAP: Parker & Vannest, 2009). These within-case effect size metrics can be used to make comparisons between pairs of phases within an SCED for a single outcome. However, in practice, many SCEDs include multiple outcomes, multiple phases (e.g., baseline, intervention, and modified intervention), or both multiple outcomes and multiple phases. In such studies, it may be useful to estimate multiple effect sizes for inclusion in a meta-analysis. This presents two challenges: a) how to account for the statistical dependency between effect size estimates that are based on a common sample and b) how to model these effect sizes. Well-developed methods are available for addressing these challenges in meta-analysis of between-groups experimental designs (Jackson, Riley, & White, 2011: Moeyaert et al., 2017: Wei & Higgins, 2013), but corresponding methods for SCEDs are lacking. Draw on the between-groups methodological literature, this study first describes methods for estimating multiple effect sizes, along with corresponding sampling variances and covariances, from SCEDs with multiple outcomes and/or multiple intervention phases. Methods are reviewed for the within-case SMD, LRR, and NAP indices (Lajeunesse, 2011: Wei & Higgins, 2013). A multi-variate, multi-level meta-analysis model is then described and applied to data from an in-progress research synthesis of SCEDs. The example illustrates how the multi-variate meta-analysis approach provides a basis for statistically comparing effects across multiple outcome domains or multiple phases and studying covariance of effects from multiple domains.

REFERENCES

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). Lawrence Erlbaum Associates, Inc.

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. Biometrics, 44(3), 837. https://doi.org/10.2307/2531595