Caveat: Brief Comments on Single-Case Effect Sizes and Meta-Analyses*

Joel R. Levin
University of Arizona

***Based, in part, on Kratochwill, T. R., & Levin, J. R. (2014). Meta- and statistical analysis of single-case intervention research data: Quantitative gifts and a wish list. *Journal of School Psychology*, *52*, 231-235.**

Possible Single-Study Effect-Size Measures

Nonoverlap measures (e.g., Parker, Vannest, & Davis, 2014)

Other "common language" (McGraw & Wong, 1992) measures:

- **Percent increase/decrease**
- **Change in pre-established classification criteria**

Add Cohen-type exclamatory modifiers, if wanted — for example, the following WRDs:

- **"Wow!" for data passing the IOTT**
- **"Respectable" for moderate-looking effects**
- **"Darn!" for weak, statistically nonsignificant effects**

Uncommon (esoteric) language measures:

- **Regression-based measures (e.g., Ross & Begeny, 2014)**
- **Percent of variance "explained" by the intervention**
- **Log response ratios (e.g., Pustejovsky, 2018)**
- **Mean phase difference in standard deviation units (*d*)**

Comments on Shadish et al.'s (2014) *d* and *g* Parametric Effect-Size Measures

- Standardization is based on combined within- and between-subjects variance.
- Why? Admirable goal: To unify effect-size calculations for the conventional "group" and single-case research literatures

Parametric Effect-Size Measures

Kudos to Shadish, Hedges, Rindskopf, Pustejovsky, and the gang!

Distinction to be made here (following Levin & Robinson, 2003) :

Single-study decision-oriented research

vs.

Multiple-study syntheses

Akin to Abelson's (1997) "lore" vs. record"

Preaching to the Choir?

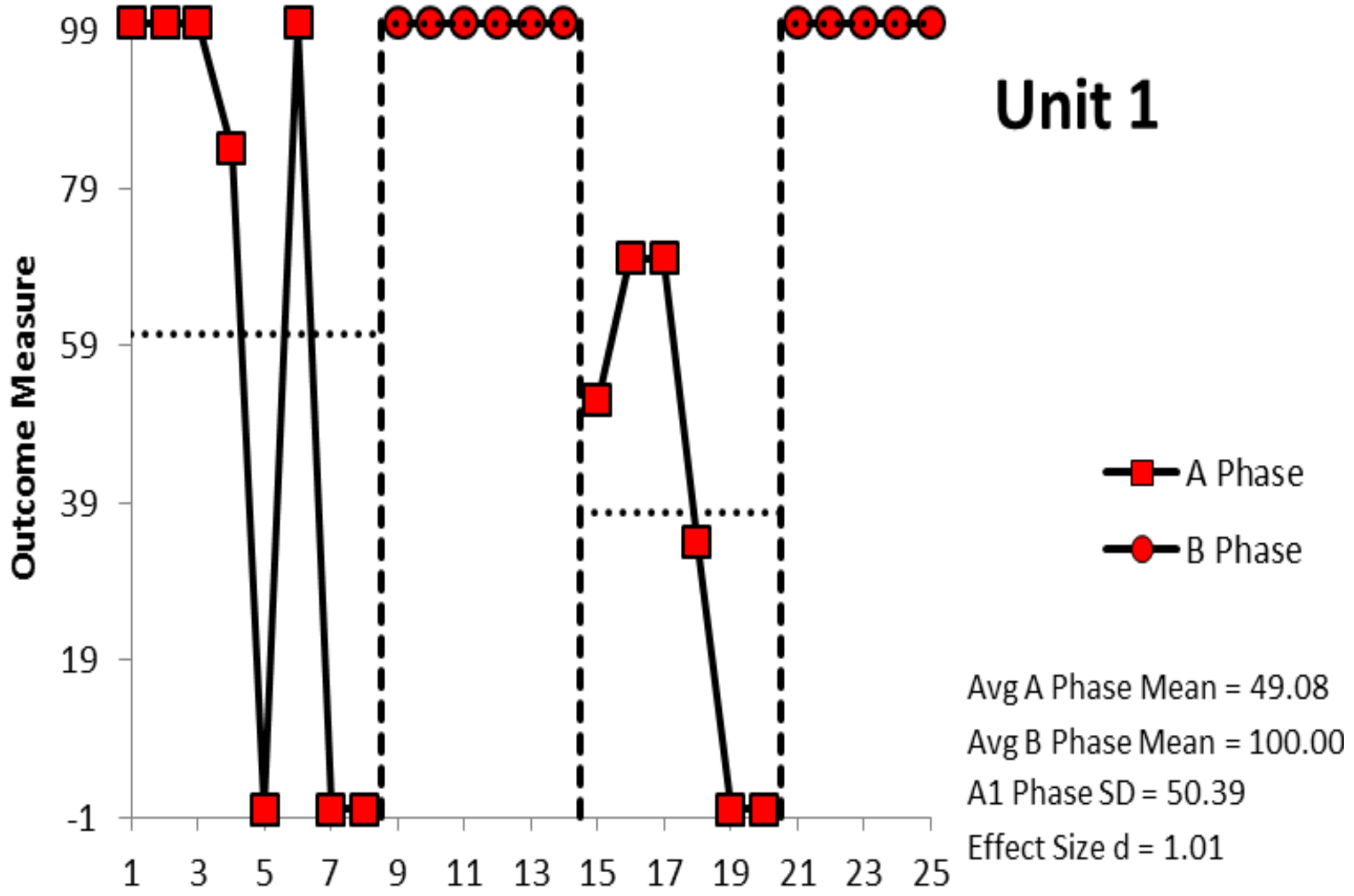
"It is not that standardizing an effect size within case is wrong. Indeed, SCD research needs effect sizes that are computed within case. The reason is that most SCD researchers are trained, and train their students, to examine each case individually, one case at a time. A study-wide summary over cases is not useful for that purpose, as useful as it may be for other purposes. SCD research needs to continue both approaches to effect size computation."

Shadish, Hedges, & Pustejovsky (2013)

Shadish et al.'s (2014) *d* and *g* Effect-Size Measures

Consider a single *N* = 1 study:

Tasky et al. (2008) Data (Participant 1)
[From Shadish et al., 2014, Appendix, pp. 552-553]

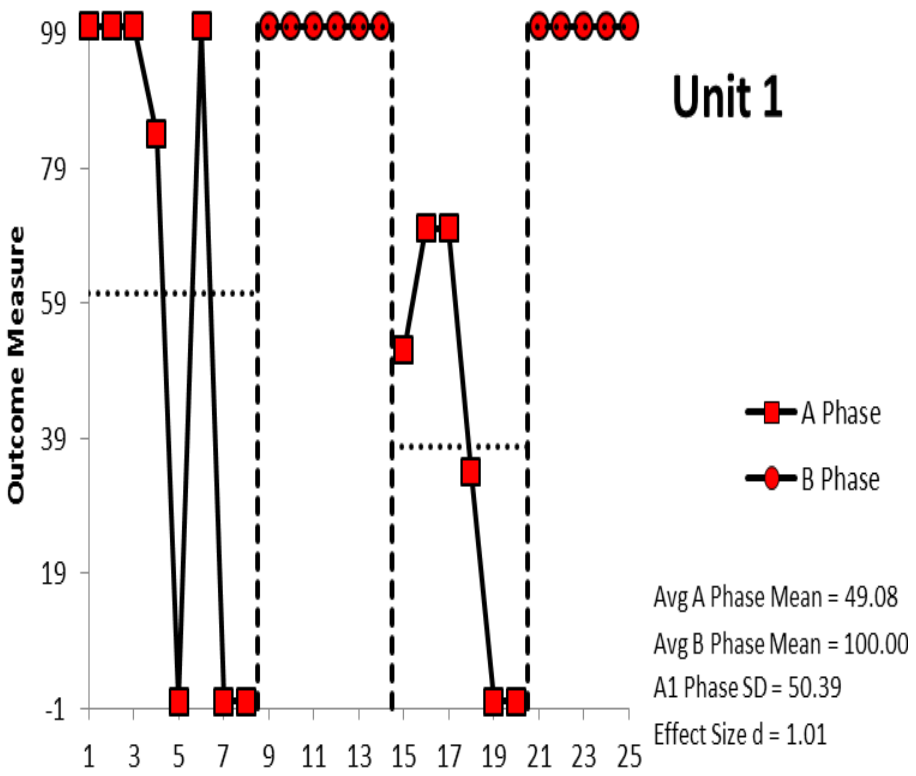


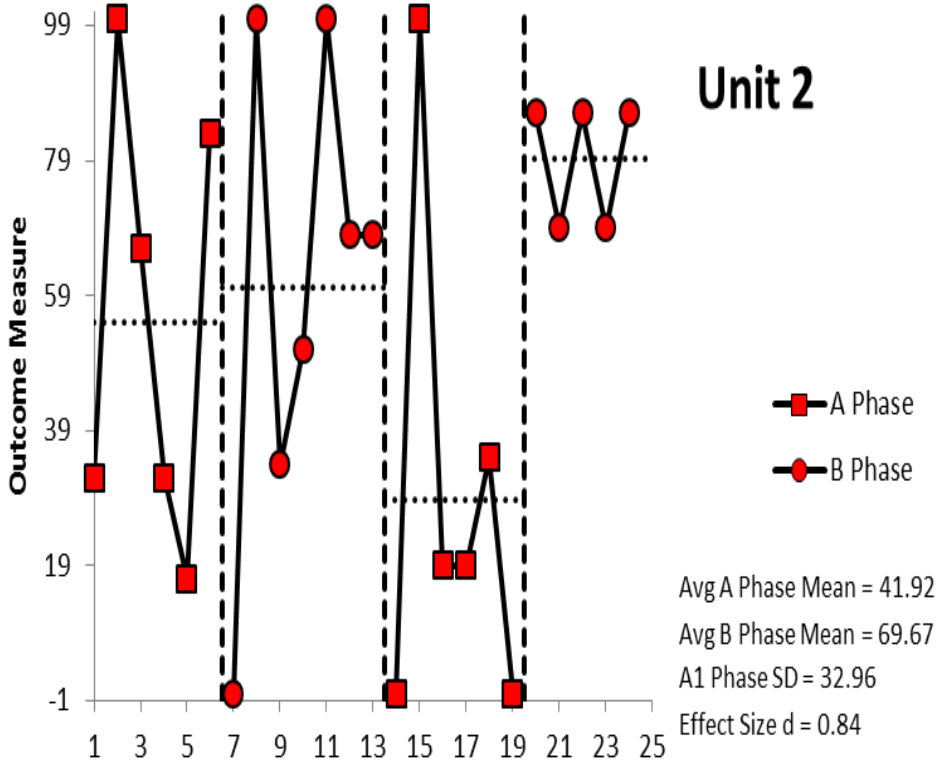
Busk & Serlin's (1992) *d* = 1.01

Shadish et al.'s (2014) *d* and *g* Effect-Size Measures

Consider a single *N* = 2 study:

Tasky et al. (2008) Data (Participants 1 and 2) [From Shadish et al., 2014, Appendix, pp. 552-553]



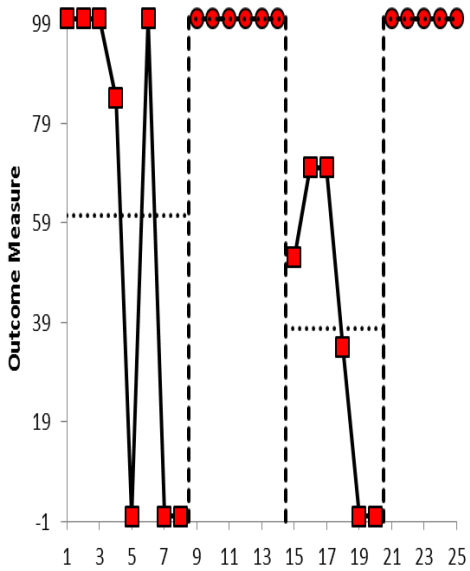


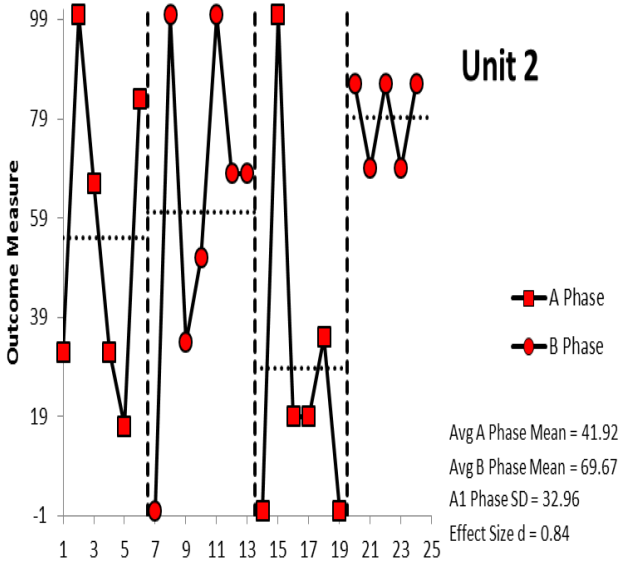
Busk & Serlin's (1992) Average *d* = 0.93

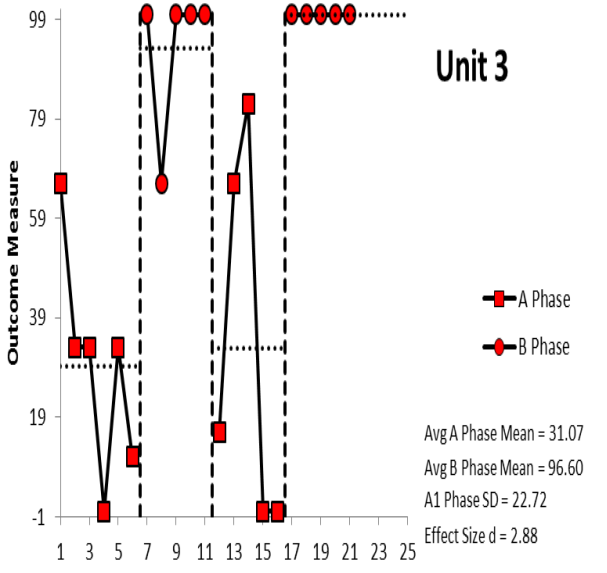
Shadish et al.'s (2014) *d* and *g* Effect-Size Measures

Consider a single *N* = 3 study:

Tasky et al. (2008) Data (Participants 1, 2, and 3) (From Shadish et al., 2014, Appendix, pp. 552-553)







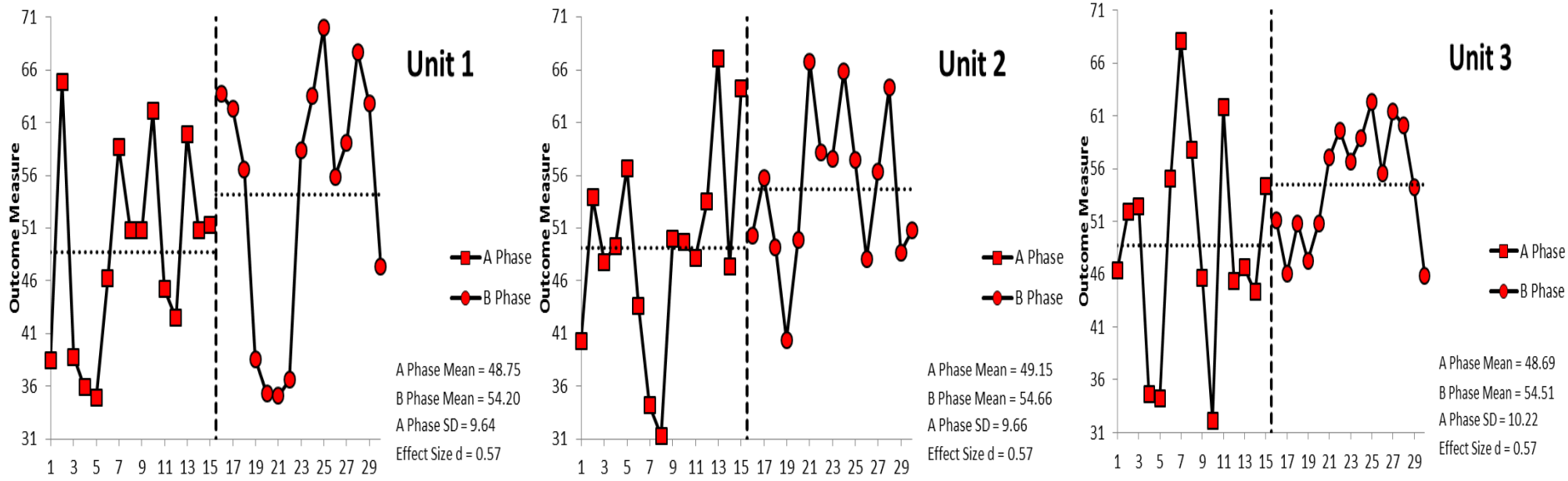
Busk & Serlin's (1992) Average *d* = 1.58
Shadish et al.'s (2014) *d* = 1.64 (*g* = 1.61)

Now, to illustrate an aspect of Shadish et al.'s *d* that might be troubling to certain single-case intervention researchers. Consider the following pair of AB design examples, each based on 3 units and 30 randomly generated outcome observations with an autocorrelation of .40. [Thanks again to John Ferron!]

Example 1:

O_A = 15, *N* (49, 100); O_B = 15, *N* (55, 100); $\Delta = (\mu_B - \mu_A)/\sigma_A$ = .60 for each unit.

d Effect-Size Visual Analysis Example 1 (A Means and Effect Sizes the "same" for all Units)



Average Effect Size *d* = 0.57

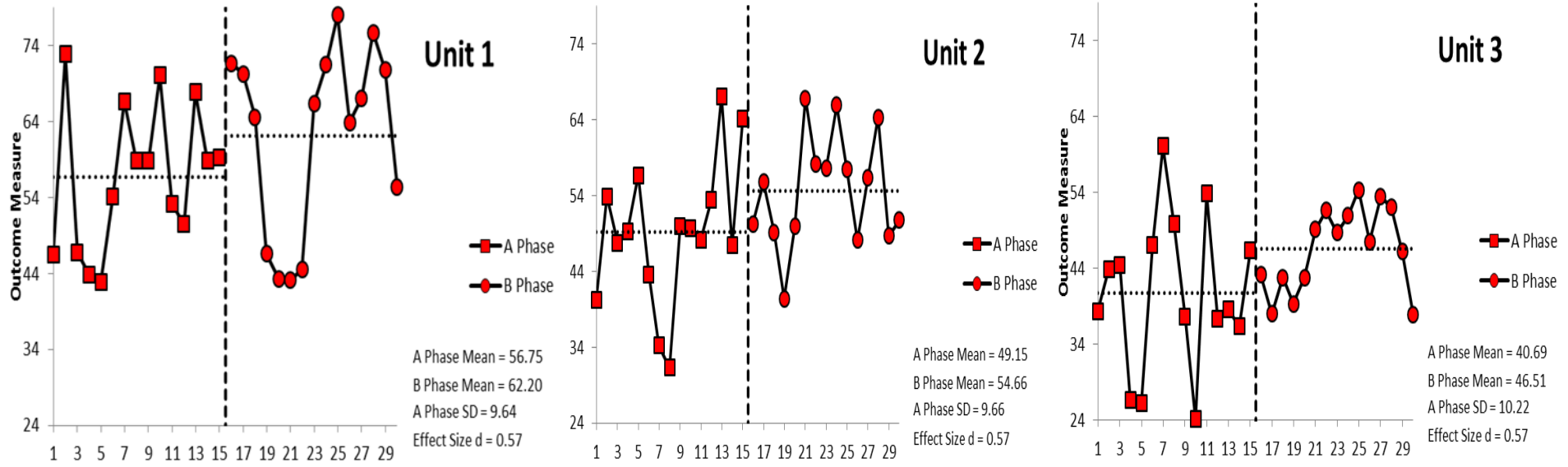
Example 2:

O_{1A} = 15, *N* (57, 100); O_{1B} = 15, *N* (63, 100); $\Delta = (\mu_B - \mu_A)/\sigma_A$ = .60.

O_{2A} = 15, *N* (49, 100); O_{2B} = 15, *N* (55, 100); $\Delta = (\mu_B - \mu_A)/\sigma_A$ = .60.

O_{3A} = 15, *N* (41, 100); O_{3B} = 15, *N* (47, 100); $\Delta = (\mu_B - \mu_A)/\sigma_A$ = .60.

d Effect-Size Visual Analysis Example 2 (A Means different for all units; Effect Sizes the same as in Example 1)



Average Effect Size *d* = 0.57

A Concluding Question to Ponder

Do we really need a mathematically unifying single-case effect-size measure such as Shadish et al.'s *d*?

1. **For meta-analysts who seek to combine effect sizes from "group" and single-case intervention studies: Yes, but maybe separate meta-analyses and associated measures should be provided for the two study types.**

2. **For meta-analysts who restrict their reviews only to single-case studies or who conduct separate analyses on single-case and "group" studies: No.**

3. **For single-case researchers with their stand-alone studies: No. If not, are there certain measures that can be recommended because they make the most conceptual and practical sense? And are there other measures that should be dismissed because they don't?**

References

Abelson, R. P. (1997). A retrospective on the significance test ban of 1999 (If there were no significance tests, they would be invented). In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.). *What if there were no significance tests?* (pp. 118-141). Mahwah, NJ: Erlbaum.

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* (pp. 187-212). Hillsdale, NJ: Erlbaum.

Levin, J. R., & Robinson, D. H. (2003). The trouble with interpreting statistically nonsignificant effect sizes in single-study investigations. *Journal of Modern Applied Statistical Methods*, *2*, 231-236.

McGraw, K. O., & Wong, S. P. (1992). A common language effect size statistic. *Psychological Bulletin, 111*, 361-365.

Parker, R. I., Vannest, K. J., & Davis, J. L. (2014) Non-overlap analysis for single-case research. In T. R. Kratochwill & J. R. Levin (eds.). (0214). *Single-case intervention research: Methodological and statistical advances* (pp. 127-151). Washington, DC: American Psychological Association.

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

Ross, S. G., & Begeny, J. C. (2014). Single-case effect size calculation: Comparing regression and non-parametric approaches across previously published reading intervention data sets. *Journal of School Psychology*, *52*, 419-431.

Shadish, W. R., Hedges, L. V., & Pustejovsky, J. E. (2013). Analysis and meta-analysis of single-case designs with a standardized mean difference statistic: A primer and applications. *Journal of School Psychology*, *52*, 123-147.

Shadish, W. R., Hedges, L. V., Pustejovsky, J. E., Boyajian, J. G., Sullivan, K. J., Andrade, A., & Barrientos, J. L. (2014). A *d*-statistic for single-case designs that is equivalent to the usual between-groups *d*-statistic. *Neuropsychological Rehabilitation*, *24*, 529-553.