Effect sizes and measurement-comparability for meta-analysis of single-case research

James E. Pustejovsky

Northwestern University
(but soon to be at University of Texas at Austin)

pusto@u.northwestern.edu

May 28, 2013

Meta-analysis of single-case research

- Summarizing results from multiple cases in a study
- Summarizing results from multiple studies
- Comparing results across variations in intervention, participant characteristics
- Means for identifying evidence-based practices/programs
- Improving external validity

MANY proposed effect size metrics

- Non-overlap metrics
  - Percentage of non-overlapping data (Scruggs, Mastropieri, & Castro, 1987)
  - Percentage of all non-overlapping data (Parker, Hagan-Burke, & Vannest, 2007)
  - Percentage exceeding median (Ma, 2006)
  - Non-overlap of all pairs (Parker & Vannest, 2009)
  - Improvement rate difference (Parker, Vannest, & Brown, 2009)
  - Tau-U (Parker, Vannest, Davis, & Sauber, 2011)

- Parametric effect sizes
  - Standardized mean difference (Busk & Serlin, 1992)
  - R-squared metrics (Center, Skiba & Casey, 1985; Allison & Gorman, 1993; Faith, Allison, Franklin, & Gorman, 1996; Beretvas & Chung, 2008)

- Hierarchical linear models
  - Van den Noortgate & Onghena (2003a, 2003b, 2007, 2008)
  - Hedges, Pustejovsky, & Shadish (2012)
  - Shadish, Kyse, & Rindskopf (2013)

Effect size desiderata

1. Should measure magnitude of treatment effect/ functional relation
   - Should not depend strongly on other procedural details of the study

2. Should include some measure of precision/uncertainty
   - Standard error
   - Confidence interval

3. Should be comparable across different methods of measuring the same construct

Shogren, et al. (2004)

The effect of choice-making as an intervention for problem behavior

- 13 single-case studies
- 32 unique cases

Measurement procedure	# Cases
Event counting	3
Continuous recording	5
Momentary time sampling	1
Interval recording	19
Other	4

Measures of percentage change

- Mean baseline reduction (Campbell, 2004; Campbell & Herzinger, 2010)
- Suppression index (Hershberger, et al., 1999; Marquis et al., 2000)

$$MBR = \left(\frac{\bar{y}_B}{\bar{y}_A} - 1 \right) \times 100\%$$

where

- $\bar{y}_A$ is the mean outcome in phase A (baseline)
- $\bar{y}_B$ is the mean outcome in phase B (treatment)

Log-Response Ratios

- Hedges, Gurevitch, & Curtis (1999)
- Natural-log scale used due to better statistical properties

$$lRR = \ln\left(\frac{\bar{y}^{B}}{\bar{y}^{A}}\right)$$

- Can be transformed into *MBR*/percentage change measure

$$MBR = \left[\exp\left(lRR\right) - 1\right] \times 100\%$$

Log-Response Ratios (continued)

- Standard error (SE)

$$\mathrm{SE}\left(lRR\right) \approx \sqrt{\frac{s_A^2}{n_A \bar{y}_A^2} + \frac{s_B^2}{n_B \bar{y}_B^2}}$$

where

- s_A^2 is the sample variance in phase A
- n_A is the number of measurements in phase A
- s_B^2 is the sample variance in phase B
- n_B is the number of measurements in phase B

Approximate 95% confidence intervals

- For the log-response ratio:

$$Lower = lRR - 2 \times \mathrm{SE}\left(lRR\right)$$

$$Upper = lRR + 2 \times \mathrm{SE}\left(lRR\right)$$

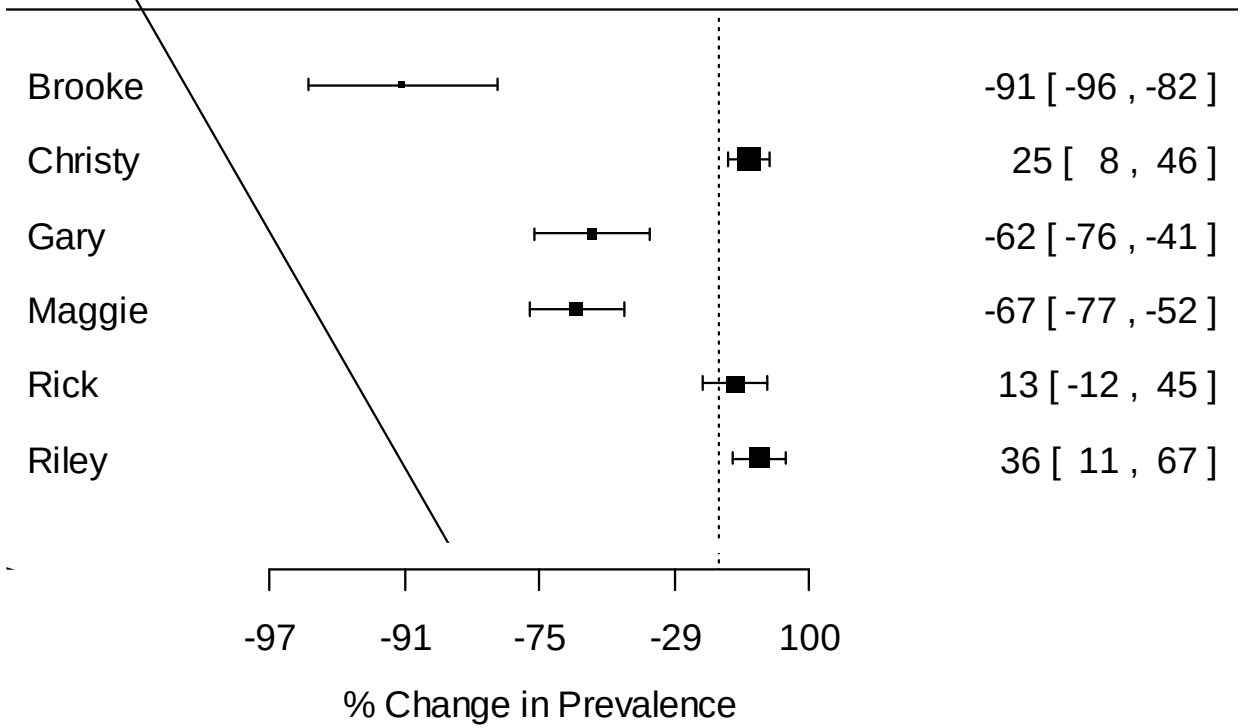
- For *MBR*/percentage change:

$$Lower = \left[\exp\left(lRR - 2 \times \mathrm{SE}\left(lRR\right)\right) - 1\right] \times 100\%$$

$$Upper = \left[\exp\left(lRR + 2 \times \mathrm{SE}\left(lRR\right)\right) - 1\right] \times 100\%$$

Romaniuk, et al. (2002)

The influence of activity choice on problem behaviors maintained by escape versus attention.[1]



1. *JABA*, *35*(4).

Measurement comparability

- When the average length (duration) of behavior does not change between phases, the response ratio based on frequency counting is equivalent to the response ratio based on continuous recording/MTS.

- For interval recording methods, further assumptions about average duration of behavior are needed in order to make direct comparisons to frequency counting or continuous recording data.

Handling interval recording data

- Scenario 1
  - average event durations are greater than some known value
- Scenario 2
  - most inter-event times are larger than the interval length
  - the average event durations are short
- Scenario 3
  - average event duration is unaffected by the treatment
  - inter-event times are exponentially distributed

Key Take-Aways

- For outcomes that are based on direct observation of behavior, consider using the log-response ratio.
- Use log-scale for calculating standard errors and for meta-analysis.
- Transform into percentage change (MBR) for easy interpretation.
- Pay careful attention to how outcomes are measured.

References

- Allison, D. B., & Gorman, B. S. (1993). Calculating effect sizes for meta-analysis: The case of the single case. *Behaviour Research and Therapy*, *31*(6), 621–31.
- 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.
- Campbell, J. M. (2004). Statistical comparison of four effect sizes for single-subject designs. *Behavior Modification*, *28*(2), 234–46.
- Campbell, J. M., & Herzinger, C. V. (2010). Statistics and single subject research methodology. In D. L. Gast (Ed.), *Single Subject Research Methodology in Behavioral Sciences* (pp. 417–450). New York, NY: Routledge.
- Center, B. A., Skiba, R. J., & Casey, A. (1985). A methodology for the quantitative synthesis of intra-subject design research. *The Journal of Special Education*, *19*(4), 387-400.
- Faith, M. S., Franklin, R. D., Allison, D. B., & Gorman, B. S. (1996). Meta-analysis of single-case research. In R. D. Franklin, D. B. Allison, & B. S. Gorman (Eds.), *Design and Analysis of Single-Case Research* (pp. 245–277). Mahwah, NJ: Lawrence Erlbaum.
- Hedges, L. V, Gurevitch, J., & Curtis, P. (1999). The meta-analysis of response ratios in experimental ecology. *Ecology*, *80*(4), 1150–1156.
- Hedges, L. V, Pustejovsky, J. E., & Shadish, W. R. (2012). A standardized mean difference effect size for single case designs. *Research Synthesis Methods*, *3*, 224–239.

References

- Hershberger, S. L., Wallace, D. D., Green, S. B., & Marquis, J. G. (1999). Meta-analysis of single-case designs. *Statistical Strategies for Small Sample Research* (pp. 109–132).
- Ma, H. H. (2006). An alternative method for quantitative synthesis of single-subject researches: Percentage of data points exceeding the median. *Behavior Modification*, *30*(5), 598.
- Marquis, J. G., Horner, R. H., Carr, E. G., et al. (2000). A meta-analysis of positive behavior support. In R. Gersten, E. P. Schiller, & S. Vaughan (Eds.), *Contemporary Special Education Research: Syntheses of the Knowledge Base on Critical Instructional Issues* (pp. 137–178). Mahwah, NJ: Lawrence Erlbaum.
- Parker, R. I., Hagan-Burke, S., & Vannest, K. (2007). Percentage of all non-overlapping data PAND: An alternative to PND. *Journal of Special Education, 40*, 194-204.
- Parker, R. I., Vannest, K., & Brown, L. (2009). The improvement rate difference for single-case research. *Exceptional Children*, *75*(2), 135–150.
- Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy*, *40*(4), 357–67.
- Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior therapy*, *42*(2), 284–299.
- Romaniuk, C., Miltenberger, R., Conyers, C., Jenner, N., Jurgens, M., & Ringenberg, C. (2002). The influence of activity choice on problem behaviors maintained by escape versus attention. *Journal of Applied Behavior Analysis*, *35*(4), 349–62.
- Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research. *Remedial and Special Education*, *8*(2), 24–43.

References

- Shadish, W. R., Kyse, E. N., & Rindskopf, D. M. (2013). Analyzing data from single-case designs using multilevel models: New applications and some agenda items for future research. *Psychological Methods*. Forthcoming.
- Shogren, K. A., Faggella-luby, M. N., Bae, S. J., & Wehmeyer, M. L. (2004). The effect of choice-making as an intervention for problem behavior. *Journal of Positive Behavior Interventions*, *6*(4), 228–237.
- Van Den Noortgate, W., & Onghena, P. (2003). Combining single-case experimental data using hierarchical linear models. *School Psychology Quarterly*, *18*(3), 325–346.
- Van den Noortgate, W., & Onghena, P. (2003). Hierarchical linear models for the quantitative integration of effect sizes in single-case research. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 1–10.
- Van den Noortgate, W., & Onghena, P. (2007). The aggregation of single-case results using hierarchical linear models. *Behavior Analyst Today*, *8*(2), 196–209.
- 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..