

JAMES PUSTEJOVSKY
UNIVERSITY OF TEXAS - AUSTIN

Effect Sizes for Single-Case Research

The research reported here was supported in part by the Institute of Education Sciences, U.S. Department of Education, through Grants R305D100046 and R305D160002. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.



EFFECT SIZES FOR SINGLE-CASE RESEARCH

- Effect sizes: *numerical indices quantifying magnitude and direction of an intervention effect (functional relationship) on a scale that is comparable across cases/studies.*
- Why does single-case research need effect size measures?
 1. Synthesizing the evidence base for an intervention/practice using meta-analysis
 2. Examining heterogeneity of effects (for whom and under what conditions is a treatment effective?)
 3. Studying the process of evidence production (meta-science)

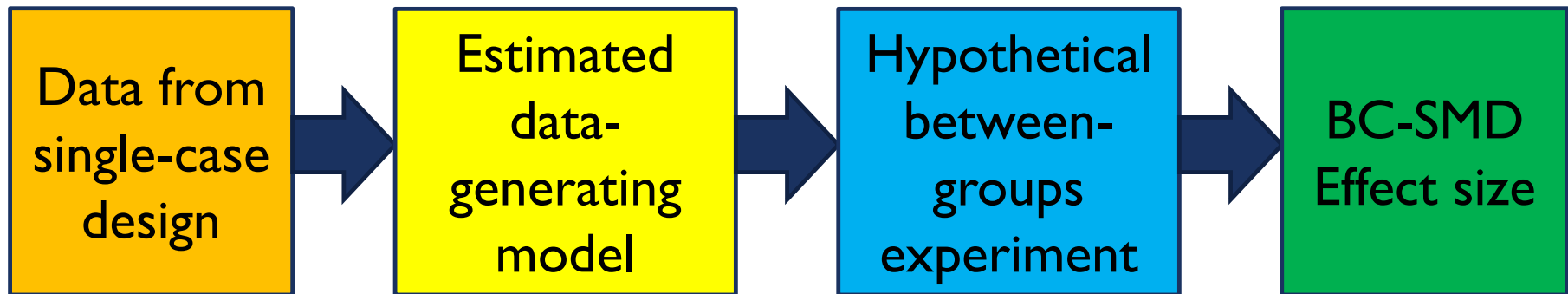
WHAT KIND OF EFFECT SIZES DO WE NEED?

- Comparability *across a set of cases/studies to be synthesized*
- Studies might use different measurement procedures to assess a common outcome construct (dependent variable)
 - E.g., continuous recording, momentary time sampling, partial interval recording of a behavior
- Studies might use different research designs to evaluate a common intervention (functional relationship)
 - E.g., multiple baseline, treatment reversal, between-case experimental designs

BETWEEN-CASE STANDARDIZED MEAN DIFFERENCES (BC-SMD)

- SMD is the most familiar effect size index in between-case experimental research
 - Difference in mean outcomes (treated – not treated)
 - Scaled by cross-sectional SD of outcome
- BC-SMDs (Hedges, Pustejovsky, & Shadish, 2012, 2013; Pustejovsky, Hedges, & Shadish, 2014; Shadish, Hedges, & Pustejovsky, 2014) estimate ***the same effect size parameter as in a between-case experimental design***, using data from a single-case design.
- Translate single-case results into terms that are familiar for between-case researchers (Shadish, Hedges, Horner, & Odom, 2015)

BETWEEN-CASE STANDARDIZED MEAN DIFFERENCES (BC-SMD)



- Methods currently available for:
 - Across-participant multiple baseline/multiple probe design
 - Treatment reversal design with replication across cases
 - Requires studies with at least 3 participants, outcomes measured on a common scale

TOOLS FOR CALCULATING BC-SMDS

- SPSS macro (Shadish & Marso, 2015)
- R package `scdhlm` available on CRAN (Pustejovsky, 2016)
- Shiny web-app now available (<https://jepusto.shinyapps.io/scdhlm/>)

The screenshot shows a web browser window displaying the Shiny application 'Between-case standardized mean difference estimator'. The browser address bar shows the URL <https://jepusto.shinyapps.io/scdhlm/>. The application interface includes a navigation bar with tabs for 'scdhlm', 'Load', 'Inspect', 'Model', and 'Effect size'. The 'Model' tab is currently selected. Below the navigation bar, there are several interactive elements:

- Estimation method:** A dropdown menu set to 'Restricted Maximum Likelihood'.
- Center session at:** A horizontal slider ranging from 1 to 10, with a blue bar and a grey knob currently positioned at 3.
- Baseline phase:** A section with a 'Type of time trend' dropdown set to 'level', and two checkboxes for 'Include fixed effect' and 'Include random effect', both of which are checked.
- Treatment phase:** A section with a 'Type of time trend' dropdown set to 'level', and two checkboxes for 'Include fixed effect' (checked) and 'Include random effect' (unchecked).

APPLICATIONS OF BC-SMDS

DEVELOPMENTAL NEUROREHABILITATION

<http://dx.doi.org/10.3109/17518423.2015.1100690>



A demonstration of how to do a meta-analysis that combines single-case designs with between-groups experiments: The effects of choice making on challenging behaviors performed by people with disabilities

Nicole A. M. Zelinsky and William Shadish

Department Psychological Sciences, School of Social Sciences, Humanities and Arts, University of California, Merced, Merced, CA, USA

ABSTRACT

Objective: This paper demonstrates how to conduct a meta-analysis that includes both between-group and single-case design (SCD) studies. The example studies whether choice-making interventions decrease challenging behaviors performed by people with disabilities. *Methods:* We used a between-case *d*-statistic to conduct a meta-analysis of 15 between-group and SCD studies of 70 people with a disability, who received a choice intervention or control. We used robust variance estimation to adjust for dependencies caused by multiple effect sizes per study, and conducted moderator, sensitivity, influence, and publication bias analyses. *Results:* The random-effects average was $d = 1.02$ (standard error of 0.168), so the 95% confidence interval (CI) suggests choice-making reduces challenging behaviors by 0.65 to 1.38 standard deviations. Studies that provided choice training produced a significantly larger intervention effect. *Conclusion:* Choice-making reduces challenging behaviors performed by people with disabilities.

KEYWORDS

Meta-analysis; challenging behaviors; robust variance estimation; choice; disabilities

APPLICATIONS OF BC-SMDS



Journal of Clinical Epidemiology 76 (2016) 82–88

Journal of
Clinical
Epidemiology

Single-case experimental design yielded an effect estimate corresponding to a randomized controlled trial

William R. Shadish^{a,†}, David M. Rindskopf^{b,*}, Jonathan G. Boyajian^c

^a*Department of Psychological Sciences, University of California Merced, 5200 N Lake Rd., Merced, CA 95343, USA*

^b*Educational Psychology, CUNY Graduate Center, New York, NY 10036-8003, USA*

^c*Public Health Group, School of Social Sciences, Humanities, and Arts, University of California Merced, Merced, CA 95343, USA*

Accepted 3 January 2016; Published online 11 April 2016

Abstract

Objectives: We reanalyzed data from a previous randomized crossover design that administered high or low doses of intravenous immunoglobulin (IgG) to 12 patients with hypogammaglobulinaemia over 12 time points, with crossover after time 6. The objective was to see if results corresponded when analyzed as a set of single-case experimental designs vs. as a usual randomized controlled trial (RCT).

Study Design and Settings: Two blinded statisticians independently analyzed results. One analyzed the RCT comparing mean outcomes of group A (high dose IgG) to group B (low dose IgG) at the usual trial end point (time 6 in this case). The other analyzed all 12 time points for the group B patients as six single-case experimental designs analyzed together in a Bayesian nonlinear framework.

Results: In the randomized trial, group A ($M = 794.93$; standard deviation (SD) = 90.48] had significantly higher serum IgG levels at time six than group B ($M = 283.89$; SD = 71.10) ($t = 10.88$; $d^2 = 10$; $P < 0.001$), yielding a mean difference of $M_D = 511.05$ [standard error (SE) = 46.98]. For the single-case experimental designs, the effect from an intrinsically nonlinear regression was also significant and comparable in size with overlapping confidence intervals: $M_D = 495.00$, SE = 54.41, and $t = 495.00/54.41 = 9.10$. Subsequent exploratory analyses indicated that how trend was modeled made a difference to these conclusions.

Conclusions: The results of single-case experimental designs accurately approximated results from an RCT, although more work is needed to understand the conditions under which this holds. © 2016 Elsevier Inc. All rights reserved.

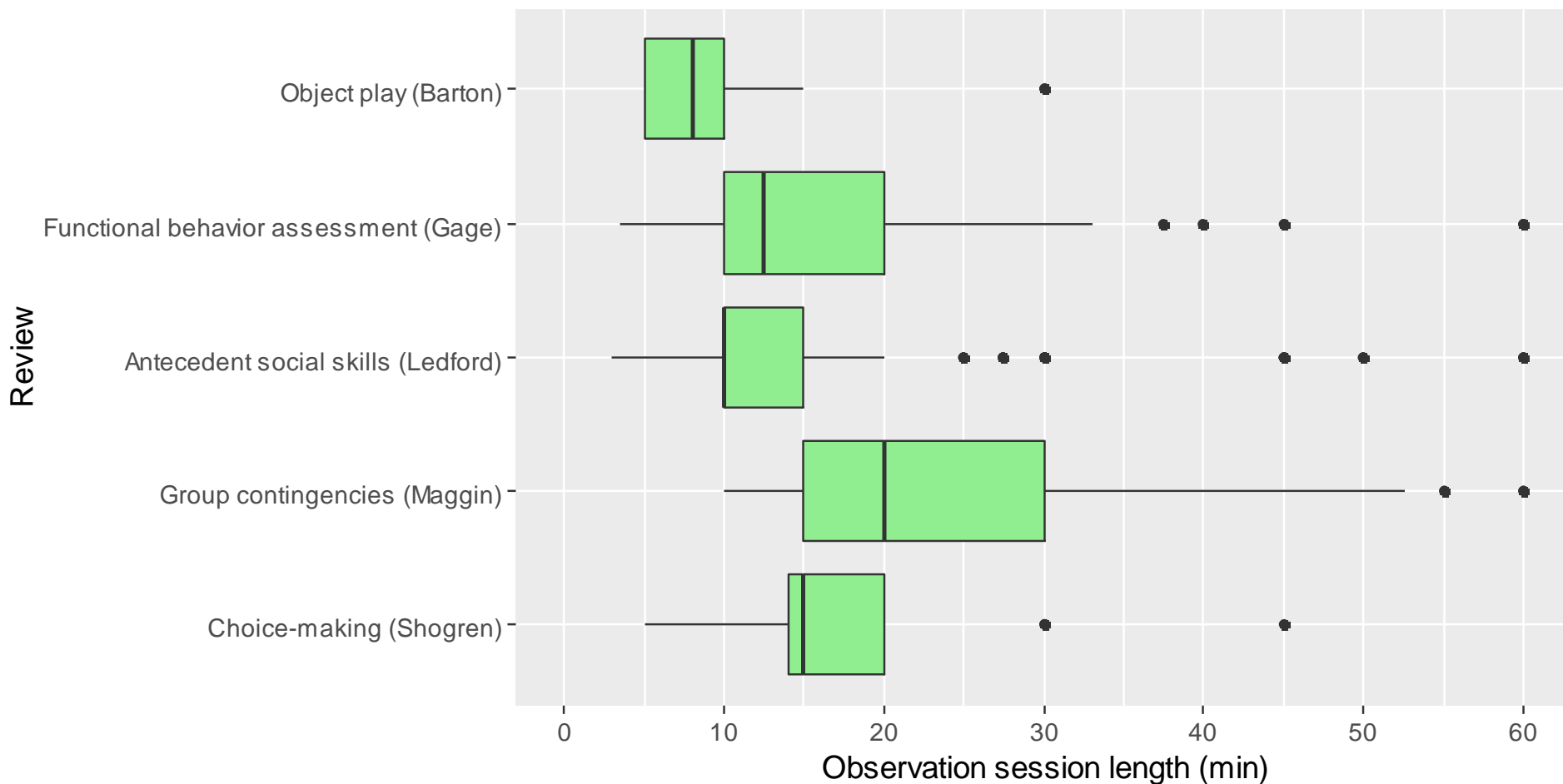
WITHIN-CASE EFFECT SIZES

- Within-case = characterize magnitude of functional relationship separately for each case in a study.
- Many different indices have been proposed
 - parametric measures (e.g., within-case standardized mean difference)
 - non-overlap measures (e.g., PND, IRD, NAP, Tau-U).
- Need for effect sizes that are on comparable scale across cases/studies that use different outcome measurement procedures.
 - Outcome recording system
 - Observation session length
 - Phase lengths

MEASUREMENT PROCEDURES IN FIVE SYSTEMATIC REVIEWS OF SINGLE-CASE RESEARCH

Procedure	Object play (Barton)	Functional behavior assessment (Gage)	Antecedent social skills (Ledford)	Group contingencies (Maggin)	Choice-making (Shogren)	Total
Event counting	11 (21%)	13 (7%)	213 (32%)	61 (36%)	8 (11%)	306 (27%)
Continuous recording		15 (8%)	49 (7%)	16 (9%)	5 (7%)	85 (7%)
Momentary time sampling		16 (8%)		34 (20%)	1 (1%)	51 (4%)
Partial interval	34 (65%)	130 (66%)	101 (15%)	55 (33%)	23 (32%)	343 (30%)
Whole interval		19 (10%)	14 (2%)		9 (13%)	42 (4%)
% correct (fixed)	3 (6%)		247 (37%)	3 (2%)		253 (22%)
% correct (variable)		2 (1%)	38 (6%)		11 (15%)	51 (4%)
Task check-list			3 (<1%)		10 (14%)	13 (1%)
Rating scale	4 (8%)	1 (1%)			4 (6%)	9 (1%)

OBSERVATION SESSION LENGTHS (MIN) IN FIVE SYSTEMATIC REVIEWS OF SINGLE-CASE RESEARCH



LOG RESPONSE RATIO

- Log response ratios quantify functional relationships *in terms of proportionate change* between phases
 - Common “informal” effect size measure (Campbell & Herzinger, 2010)
 - Precedents in systematic reviews of single-case research (Campbell, 2003; Kahng, Iwata, & Lewin, 2002; Marquis et al., 2000)

- Letting μ_A, μ_B denote mean levels of outcome in phases A and B, LRR is

$$\psi = \ln(\mu_B / \mu_A) = \ln(\mu_B) - \ln(\mu_A)$$

- Useful for behavioral outcomes measured by direct observation (Pustejovsky, 2015)
 - Magnitude remains stable when outcomes are measured using different procedures.
 - Under certain conditions, also comparable across dimensional constructs.

FUTURE DIRECTIONS

- Further development of within-case log response ratios
 - to accommodate linear/non-linear time trends
 - Estimation in the presence of auto-correlated repeated measurements
- Choice of effect size metric depends on characteristics of set of studies to be synthesized.
- Publication/outcome reporting bias represents a major outstanding challenge for synthesis of single-case research.

*A SURVEY OF PUBLICATION PRACTICES OF SINGLE-CASE
DESIGN RESEARCHERS WHEN TREATMENTS HAVE SMALL
OR LARGE EFFECTS*

WILLIAM R. SHADISH, NICOLE A. M. ZELINSKY, AND JACK L. VEVEA

UNIVERSITY OF CALIFORNIA, MERCED

AND

THOMAS R. KRATOCHWILL

UNIVERSITY OF WISCONSIN, MADISON

The published literature often underrepresents studies that do not find evidence for a treatment effect; this is often called *publication bias*. Literature reviews that fail to include such studies may overestimate the size of an effect. Only a few studies have examined publication bias in single-case design (SCD) research, but those studies suggest that publication bias may occur. This study surveyed SCD researchers about publication preferences in response to simulated SCD results that show a range of small to large effects. Results suggest that SCD researchers are more likely to submit manuscripts that show large effects for publication and are more likely to recommend acceptance of manuscripts that show large effects when they act as a reviewer. A non-trivial minority of SCD researchers (4% to 15%) would drop 1 or 2 cases from the study if the effect size is small and then submit for publication. This article ends with a discussion of implications for publication practices in SCD research.

Key words: single-case design, publication bias, effect size

REFERENCES

- Campbell, J. M. (2003). Efficacy of behavioral interventions for reducing problem behavior in persons with autism: a quantitative synthesis of single-subject research. *Research in Developmental Disabilities, 24*(2), 120–138. doi:10.1016/S0891-4222(03)00014-3
- 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.
- 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. DOI: 10.1002/jrsm.1052
- Hedges, L. V, Pustejovsky, J. E., & Shadish, W. R. (2013). A standardized mean difference effect size for multiple baseline designs across individuals. *Research Synthesis Methods, 4*(4), 324–341. DOI: 10.1002/jrsm.1086
- Kahng, S., Iwata, B. a., & Lewin, A. B. (2002). Behavioral treatment of self-injury, 1964 to 2000. *American Journal of Mental Retardation : AJMR, 107*(3), 212–221. doi:10.1352/0895-8017(2002)107<0212:BTOSIT>2.0.CO;2
- Marquis, J. G., Horner, R. H., Carr, E. G., Turnbull, A. P., Thompson, M., Behrens, G. A., ... Doolabh, A. (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 Associates.
- Pustejovsky, J. E., Hedges, L. V, & Shadish, W. R. (2014). Design-comparable effect sizes in multiple baseline designs: A general modeling framework. *Journal of Educational and Behavioral Statistics, 39*(5), 368–393. DOI: 10.3102/1076998614547577
- Pustejovsky, J. E. (2015). Measurement-comparable effect sizes for single-case studies of free-operant behavior. *Psychological Methods, 20*(3), 342–359. doi:10.1037/met0000019
- Shadish, W. R., Hedges, L. V, & Pustejovsky, J. E. (2014). Analysis and meta-analysis of single-case designs with a standardized mean difference statistic: A primer and applications. *Journal of School Psychology, 52*(2), 123–147. DOI: 10.1016/j.jsp.2013.11.005
- Shadish, W. R., Hedges, L. V, Horner, R. H., & Odom, S. L. (2015). *The role of between-case effect size in conducting, interpreting, and summarizing single-case research*. Washington, DC. Retrieved from <http://ies.ed.gov/ncser/pubs/2015002/>
- Shadish, W. R., Rindskopf, D. M., & Boyajian, J. G. (2016). Single-case experimental design yielded an effect estimate corresponding to a randomized controlled trial. *Journal of Clinical Epidemiology, 76*, 82–88. DOI: 10.1016/j.jclinepi.2016.01.035
- Zelinsky, N. A. M., & Shadish, W. R. (2016). A demonstration of how to do a meta-analysis that combines single-case designs with between-groups experiments: The effects of choice making on challenging behaviors performed by people with disabilities. *Developmental Neurorehabilitation, 8423*(February), 1–13. DOI: 10.3109/17518423.2015.1100690

SYSTEMATIC REVIEWS

- Barton, E. E., Sweeney, E., & Gossett, S. (2016). A review of object play interventions for young children with disabilities. Manuscript in preparation.
- Gage, N. A., Lewis, T. J., & Stichter, J. P. (2012). Functional behavioral assessment-based interventions for students with or at risk for emotional and/or behavioral disorders in school: A hierarchical linear modeling meta-analysis. *Behavioral Disorders, 37*(2), 55–77.
- Ledford, J. R., King, S., Harbin, E. R., & Zimmerman, K. N. (2016). Antecedent Social Skills Interventions for Individuals With ASD: What Works, for Whom, and Under What Conditions? *Focus on Autism and Other Developmental Disabilities, 1–11*. DOI: 10.1177/1088357616634024
- Maggin, D. M., Johnson, A. H., Pustejovsky, J. E., Shadish, W. R., & Chafouleas, S. M. (2016). A meta-analysis of school-based group contingency interventions for students with challenging behavior: An update. Manuscript under submission.
- 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. article. DOI: 10.1177/10983007040060040401