Design-Comparable Effect Size for Single-Case Design Studies

Mariola Moeyaert, Ph.D.
Associate Professor
The State University of New York at Albany

Daniel M. Swan, Ph.D.
Postdoctoral Scholar
University of Oregon

Emily Tanner-Smith, Ph.D.
Thomson Professor, University of Oregon
What Works Clearinghouse

Presenters

Mariola Moeyaert, Ph.D.
The State University of New York
at Albany

Daniel M. Swan, Ph.D.
University of Oregon

Learning goals for this webinar

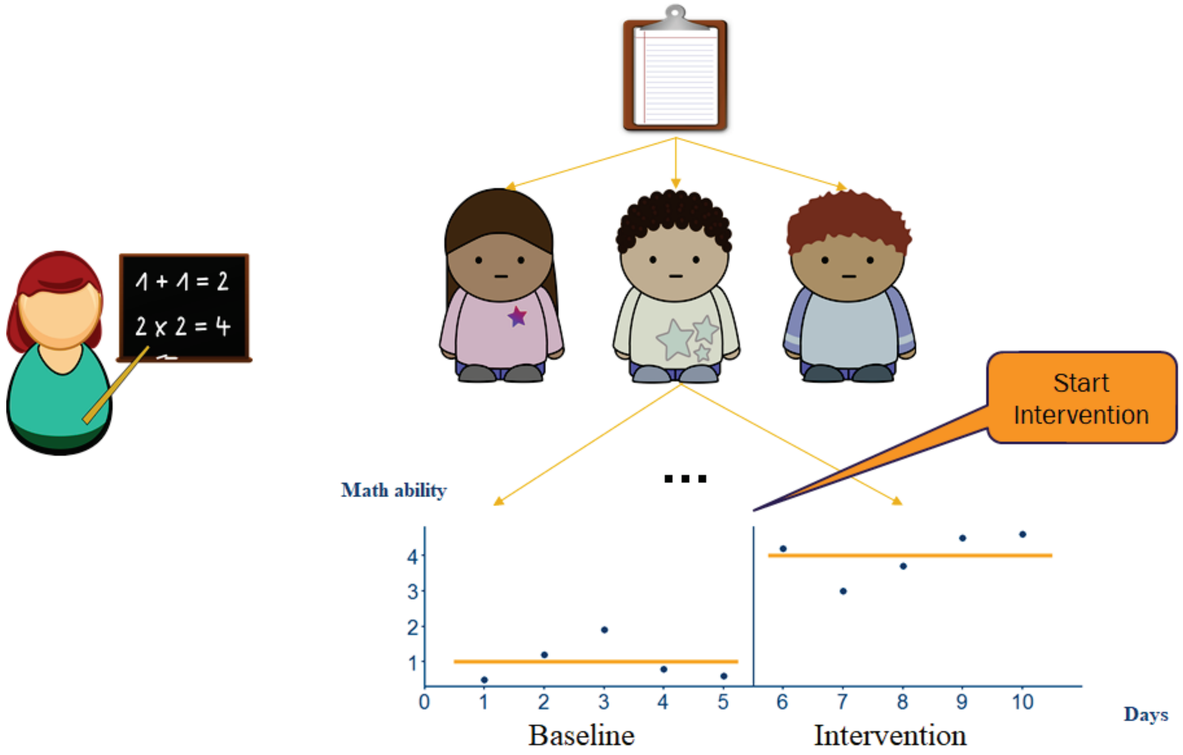
- After this webinar, you will be able to
  - Understand the difference in quantifying intervention effectiveness between Version 4.0 and Version 4.1 of the What Works Clearinghouse (WWC) Standards.
  - Quantify intervention effectiveness in single-case design (SCD) studies using the design-comparable effect size (D-CES).
  - Understand the logic behind the D-CES and its underlying assumptions.
  - Understand the advantages and disadvantages of the D-CES.
  - Use a point-and-click shiny application to calculate the D-CES.

Part 1: Introduction to SCDs and D-CES

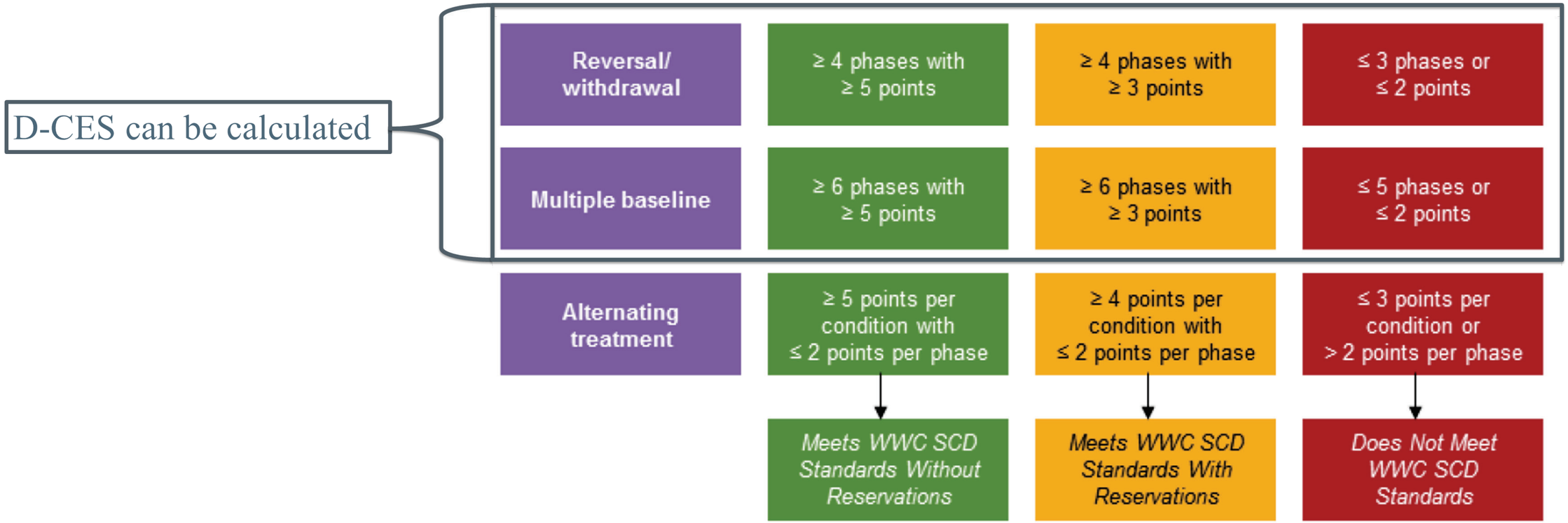
1. Introduction to SCD types
2. Approaches to quantifying intervention effects
3. Examining evidence for intervention effects:
   What Works Clearinghouse Version 4.0 versus Version 4.1
4. Introduction to D-CES

1. Introduction to SCD types

- SCDs are experiments in which one unit is observed repeatedly during a certain period of time under different levels of at least one manipulated variable.
- SCDs can demonstrate a causal effect.
- SCDs involve repeated, systematic measurement of a dependent variable before, during, and after the active manipulation of an independent variable.



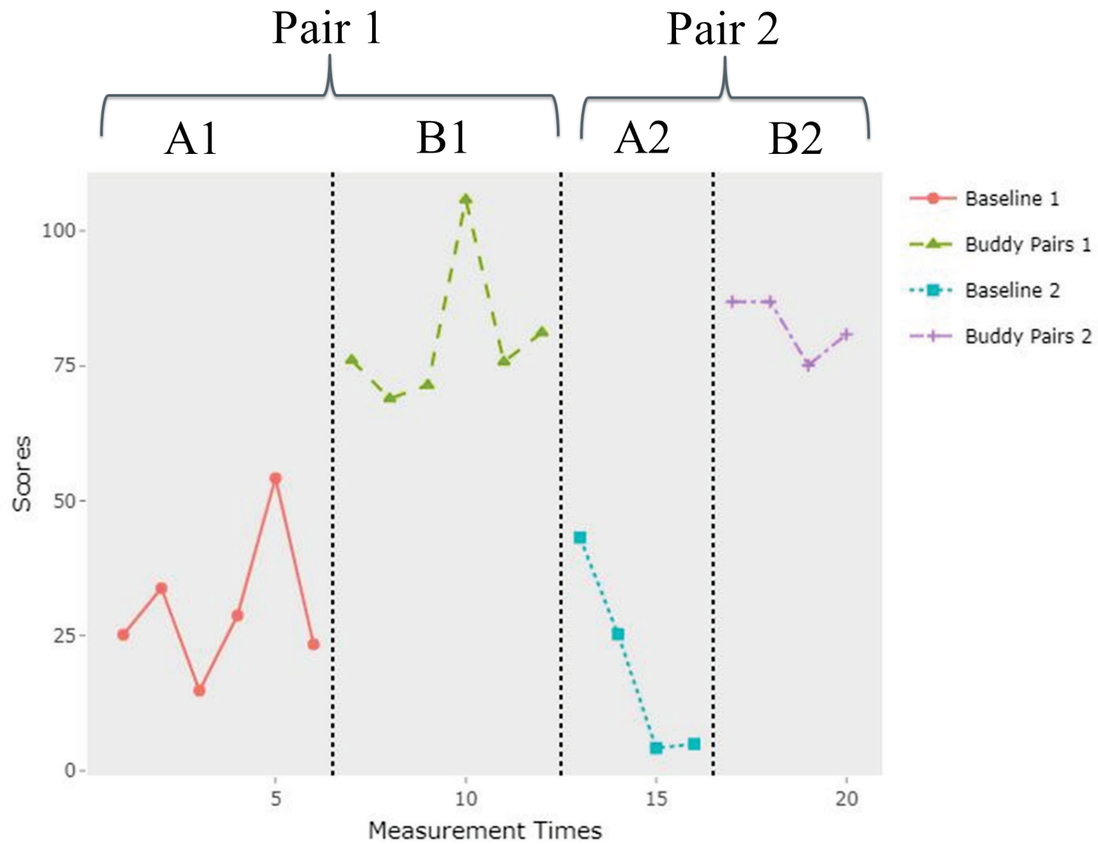
1. Introduction to SCD types



What Works Clearinghouse, Institute of Education Sciences, U.S. Department of Education. (2020). *WWC version 4.1 standards and procedures handbooks.* Retrieved October 2, 2020, from https://ies.ed.gov/ncee/wwc/Handbooks

1. Introduction to SCD types: Treatment reversal (AB^k)

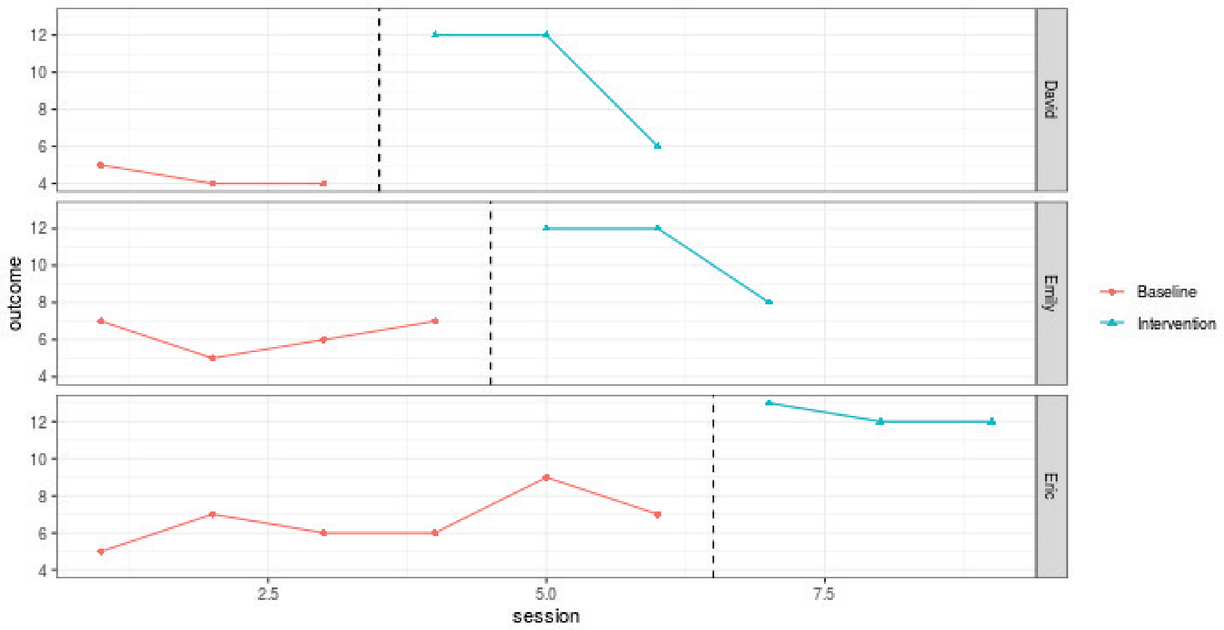
- Outcome: Percentage of appropriate social skills
- Intervention: Peer tutor training program
- Sample: Children with autism
- Three potential demonstrations for intervention effectiveness
- Data points per phase:
  - Baseline 1 and Intervention 1 → five data points
  - Baseline 2 and Intervention 2 → four data points



Laushey, K. M., & Heflin, L. J. (2000). Enhancing social skills of kindergarten children with autism through the training of multiple peers as tutors. *Journal of Autism Developmental Disorders, 30*, 183–193. Retrieved October 2, 2020, from https://doi.org/10.1023/A:1005558101038

1. Introduction to SCD types: Multiple baseline design (MB)

- Outcome: Quality of writing points
- Intervention: Self-regulated strategy development
- Sample: Three students with learning disorders
- Data points per phase:
  - Baseline → three/five data points
  - Intervention → three data points
- A least three potential demonstrations for intervention effectiveness



Saddler, B., Asaro-Saddler, K, Moeyaert, M., & Slichko, J. (2019). Teaching summary writing to students with learning disabilities via strategy instruction. *Reading & Writing Quarterly, 35*, 572–586. Retrieved October 2, 2020, from doi.org/10.1080/10573569.2019.1600085

2. Approaches to quantifying intervention effects

2.1 Within-case quantification

- Non-overlap measures.
- Log response ratio.
- Within-case standardized mean difference.
- Regression-based measures.

2.2 Across-case quantification

- Mean/median/range of the within-case quantification.
- Hierarchical linear modeling.
- **Design-comparable effect size.**

3. Examining evidence for intervention effects

WWC Version 4.0

- Intervention effects were synthesized if studies met the 5-3-20 rule:
  - At least five studies met WWC SCD standards with or without reservations, the studies were conducted by at least three different research teams with no overlapping authorship at three different institutions, and the combined number of cases was at least 20.
- Review teams tallied and computed the proportion of SCD intervention effects that were positive or negative.
- SCD effects were not synthesized with effects from group design studies.

3. Examining evidence for intervention effects

WWC Version 4.1

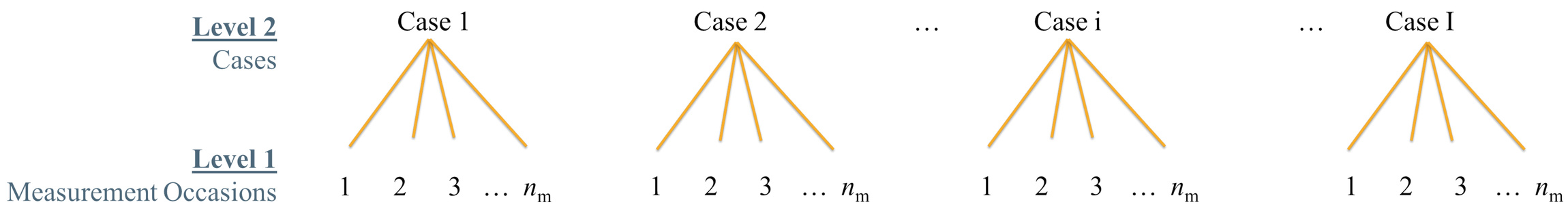
- Intervention effects are synthesized using the D-CES.
- The D-CES can be synthesized with effects from group-design studies.
- The D-CES are combined across studies using a fixed-effects meta-analysis. Fixed-effects meta-analysis involves computing a weighted average effect size. Studies are weighted by the inverse of the sampling variance of their effect sizes.

4. Introduction to D-CES

- D-CES estimates the same parameter from what may be different designs (Pustejovsky, Hedges, & Shadish, 2014).
- Study requirements for D-CES:
  - Design: Treatment reversal (AB^k), and multiple baseline (MB)/multiple probe
- The outcome is measured on a continuous scale that is common across cases.

4. Introduction to D-CES

- Sample size: Three or more cases
- Use of hierarchical linear modeling
- Estimation procedure: Moment estimation techniques (restrictive assumptions) or restricted maximum likelihood estimation (more general and flexible)



4. D-CES general framework

D-CES can be estimated using two-level hierarchical modeling:

Original D-CES using basic two-level model

Level 1: $\boldsymbol{Y_{it} = \beta_{0i} + \beta_{1i} Intervention_{it} + e_{it}}$

β_{1t} indicates the individual-specific intervention effect.

Level 2:

$$\boldsymbol{\beta_{0i} = \theta_0 + u_i}$$
$$\boldsymbol{\beta_{1i} = \theta_1}$$

$\boldsymbol{\theta_1}$ indicates the unstandardized intervention effect across the *i* cases.

D-CES

$$\boldsymbol{\delta = \frac{\theta_1}{\sqrt{\sigma^2 + \tau^2}}}$$

Numerator: Unstandardized intervention effect

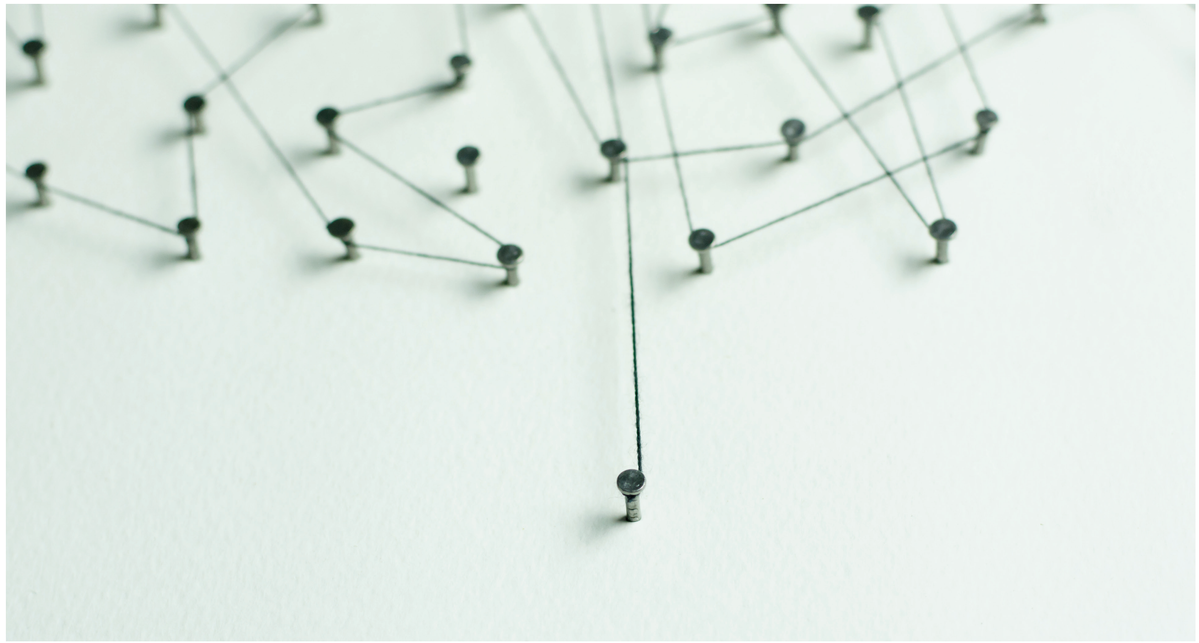
Denominator: The standard deviation of the outcome, including both within- and between-case variation

- *i* indicates the case (*i* = 1 to *I*), and case i is measured for a total of n_m measurement occasions ($t = 1, \ldots n_m$).
- Y_{it} indicates the outcome for case *i* at measurement occasion *t*.
- $Intervention_{it}$ is a dummy variable indicating whether Y_{it} is obtained during the baseline or the intervention phase.
- $e_{it} \sim N\ (0, \sigma^2)$, and the errors for case *i* follow an *AR(1)* process; $u_i \sim N\ (0, \tau^2)$.

4. D-CES general framework

- Basic D-CES assume changes in level and constant intervention effects across cases.

- More general/complex D-CES
  - Allow intervention effect to vary across cases.
  - Can include linear/polynomial time trends.



4. D-CES general framework

- If the individual-level model includes time trends for the baseline or intervention phase, one must make assumptions about how the time trends vary across cases:

  For example, Level 1: $Y_{it} = \beta_{0i} + \boldsymbol{\beta_{1i}} \times \boldsymbol{Time_{it}} + \beta_{2i} Intervention_{it} + (\boldsymbol{Time'_{it}} \times \boldsymbol{\beta_{3i}} Intervention_{it}) + e_{it}$

- If time trends in either phase vary across cases, then the total variation in the outcome is not constant → consequences for the denominator of the D-CES.
  - See Pustejovsky and colleagues (2014).

$$\delta = \frac{\boldsymbol{\theta_1}}{\sqrt{\sigma^2 + \tau^2}}$$

4. D-CES general framework

- D-CES for both AB^k and MB can be corrected for small-sample-size bias,

$$J(v) = 1 - \frac{3}{4v - 1}$$

where v is an estimated degrees of freedom (this will be somewhere between the number of cases and the total number of time points; computation of v is different for AB^k and MB).

Bias-corrected effect size:

$$g = J(v) \times \hat{\delta}$$

Part 2: Appropriate use and application of D-CES

1. Appropriately using D-CES
2. Using scdhlm

Appropriately using D-CES

- The D-CES is appropriate only for designs in which comparisons can be made across participants.
  - It is inappropriate for designs with multiple baselines across contexts or behaviors.
- The outcomes used to estimate a single D-CES should all be the same measure.

Appropriately using D-CES

- Expanding the hierarchical model
  - When there are only a few observations per phase, apparent trends might be variability.
  - When there are only a few cases, additional model parameters are difficult to estimate.

Using scdhlm

- The R package scdhlm was written by James E. Pustejovsky (Pustejovsky, 2016) for estimating the D-CES, which he also calls the "between-case standardized mean difference."
  - App website: https://jepusto.shinyapps.io/scdhlm/

- The app can be run locally on your computer by installing R and installing the scdhlm package.

- It is also possible to estimate the effect size using R code instead of the app, if you are familiar with using R.

Data organization

For the scdhlm app, data need to be organized in a “long” format, with columns including the following, at a minimum:

- The case identifier.
- The phase identifier.
- The observation session number (whole numbers are best).
- The outcome.

	A	B	C	D	E
1	**Participant Name or ID**	**Behavior Outcome Context or Setting**	**Phase Name**	**Session Number (x-axis)**	**Outcome Value (y-axis)**
2	Gizelle	Task Engagement	Baseline	1	11.9920291
3	Gizelle	Task Engagement	Baseline	2	20.03336655
4	Gizelle	Task Engagement	Baseline	3	9.936649906
5	Gizelle	Task Engagement	Baseline	4	21.98753389
6	Gizelle	Task Engagement	Intervention	5	100.1008411
7	Gizelle	Task Engagement	Intervention	6	100.1234562
8	Gizelle	Task Engagement	Intervention	7	100.1457006
9	Gizelle	Task Engagement	Intervention	8	100.1683157
10	Gizelle	Task Engagement	Intervention	9	100.1909308
11	Gizelle	Task Engagement	Baseline	10	7.038580068

About scdhlm

Between-case standardized mean difference estimator

scdhlm | Load | Inspect | Model | Effect size

- About
- Accessing scdhlm
- References
- Example data

scdhlm

Version 0.4.2.9000

Designed and maintained by James E. Pustejovsky

- pustejovsky@wisc.edu
- https://jepusto.com

Contributions from

- Bethany Hamilton
- Man Chen

Source code available on Github

Your comments, suggestions, and feedback are welcome.

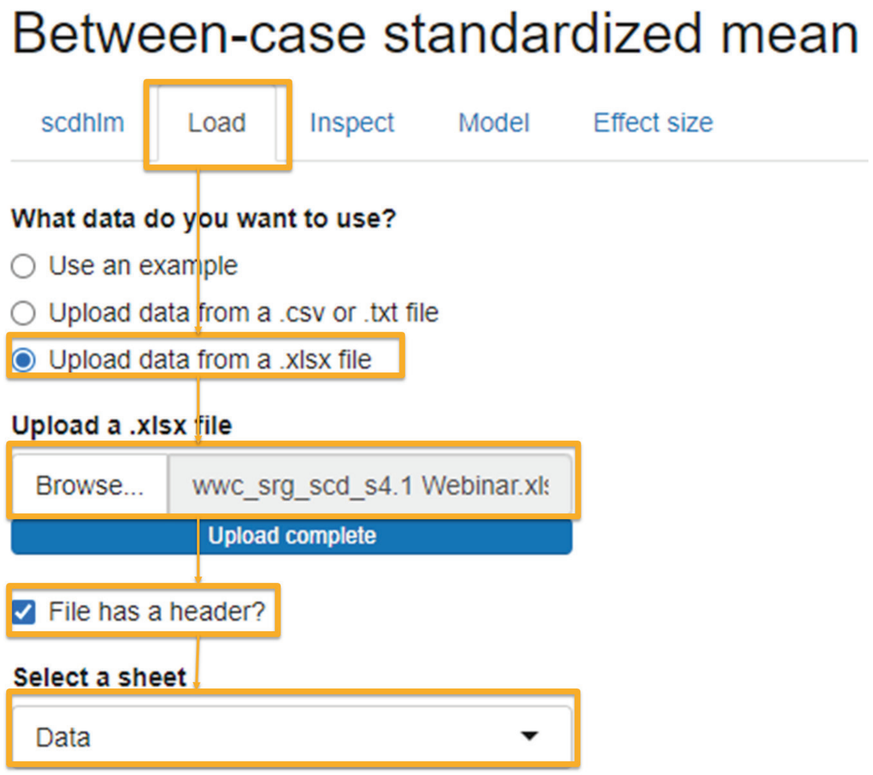
Suggested citation

Pustejovsky, James E. (2020). scdhlm: A web-based calculator for between-case standardized mean differences (Version 0.4.2.9000) \[Web application\]. Retrieved from: https://jepusto.shinyapps.io/scdhlm

Tutorial paper

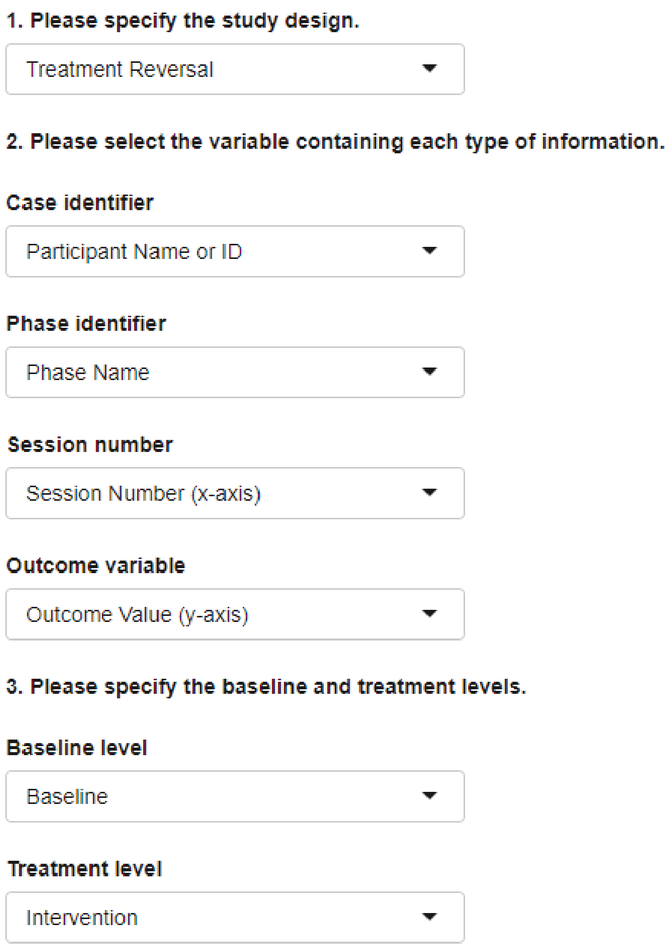
Valentine, J. C., Tanner-Smith, E. E., & Pustejovsky, J. E. (2016). Between-case standardized mean difference effect sizes for single-case designs: A primer and tutorial using the scdhlm web application. Oslo, Norway The Campbell Collaboration. DOI: 10.4073/cmpn.2016.3

Load data



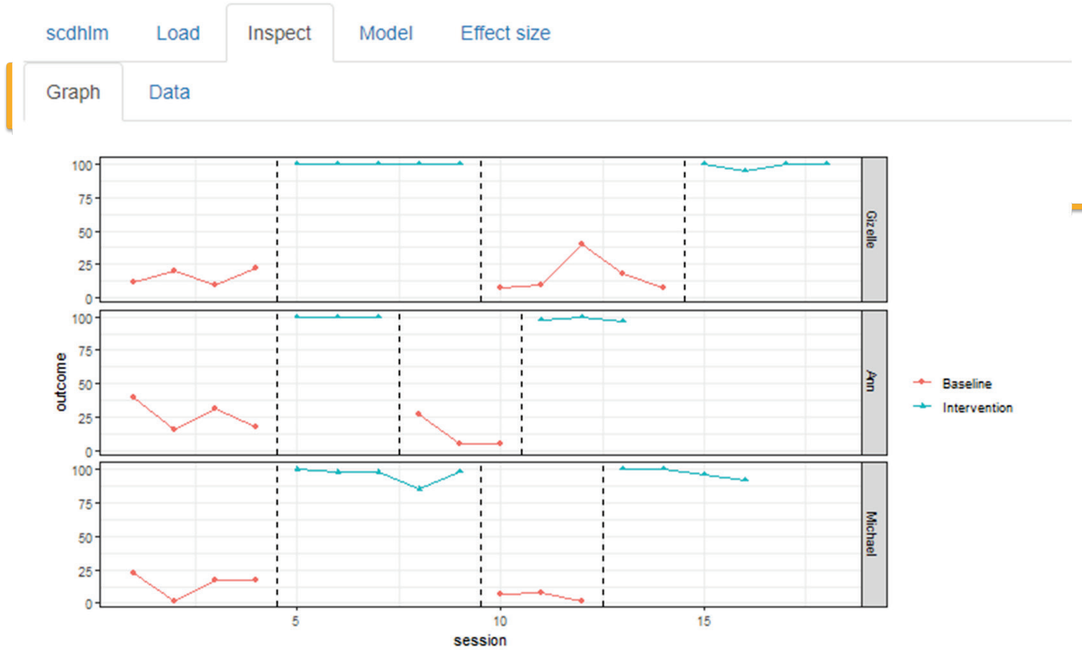
- Select the “Load” tab along the top.
- If you are going to estimate the effect size from the Study Review Guide, select the “Upload data from a .xlsx file” radio button.
- Browse to the Study Review Guide.
- Leave “File has header?” selected.
- Select the “Data” sheet.

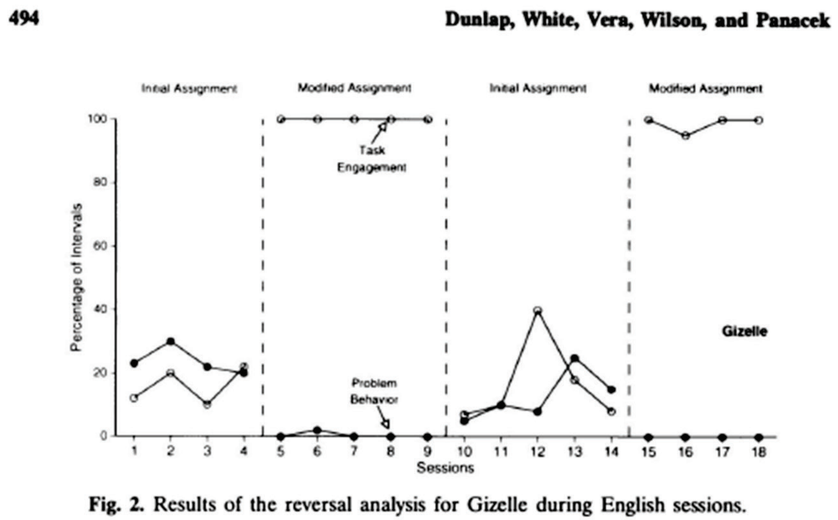
Specify variables



- Select "Treatment Reversal" under study design. If using a multiple baseline or multiple probe design, select "Multiple Baseline."
- Select the variables that correspond to the case ("Participant Name or ID"), the phase identifier ("Phase Name"), the session number ("Session Number (x-axis)"), and the outcome variable ("Outcome Value (y-axis)").
- Select the values from the "Phase Name" variable that correspond to the baseline and treatment levels.

Inspect the data





- The app will replot the data on the “Inspect” tab.
- Take the time to ensure that the replotted data look similar to the plots from which you extracted them.

Inspect the data

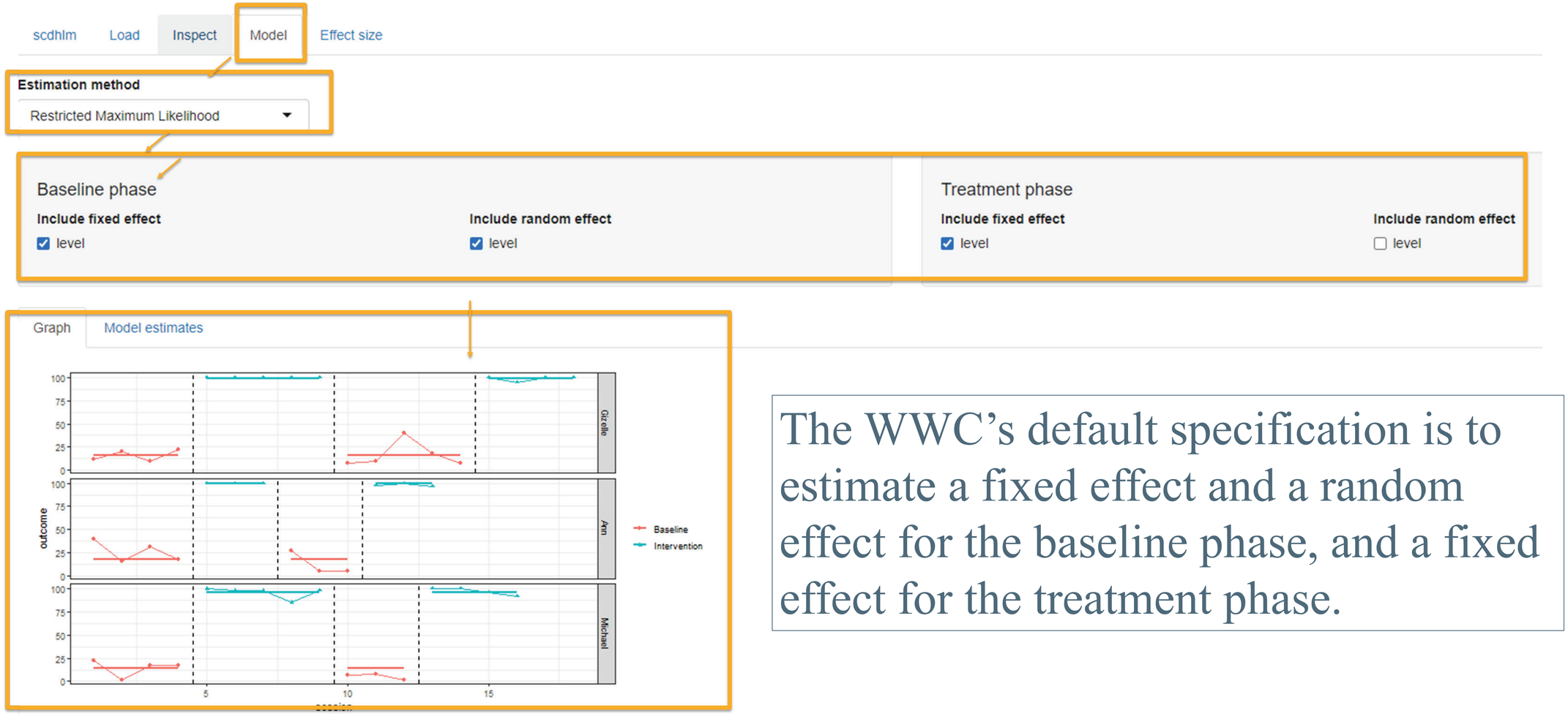
scdhlm | Load | Inspect | Model | Effect size

Graph | Data

case	session	phase	outcome	trt	phase_pair
Gizelle	1	Baseline	11.99	0.00	1.00
Gizelle	2	Baseline	20.03	0.00	1.00
Gizelle	3	Baseline	9.94	0.00	1.00
Gizelle	4	Baseline	21.99	0.00	1.00
Gizelle	5	Intervention	100.10	1.00	1.00
Gizelle	6	Intervention	100.12	1.00	1.00
Gizelle	7	Intervention	100.15	1.00	1.00
Gizelle	8	Intervention	100.17	1.00	1.00
Gizelle	9	Intervention	100.19	1.00	1.00
Gizelle	10	Baseline	7.04	0.00	2.00
Gizelle	11	Baseline	10.12	0.00	2.00
Gizelle	12	Baseline	39.92	0.00	2.00

- The four values you specified for the case, session number, phase identifier, and outcome variable are found in the first four columns.
- The last two columns are values that the app will use internally to estimate the multilevel model from which the app extracts parameters to estimate the D-CES.

Estimate the model



The WWC's default specification is to estimate a fixed effect and a random effect for the baseline phase, and a fixed effect for the treatment phase.

Record the design-comparable effect size

The “Effect size” tab contains the important information you need to record.

scdhlm | Load | Inspect | Model | Effect size

Effect size estimates and auxilliary information

CI coverage level (%) 95

BC-SMD estimate	Std. Error	95% CI (lower)	95% CI (upper)	Degrees of freedom	Auto-correlation	Intra-class correlation
9.9887	1.2201	7.5181	12.4593	37.6700	-0.0449	0.0802

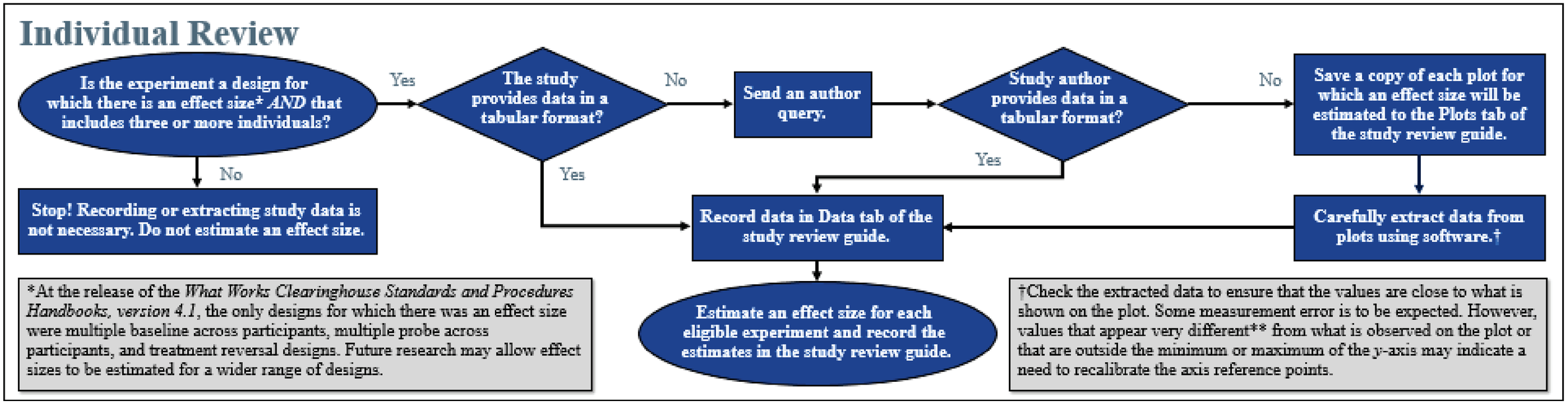
Record the design-comparable effect size

The “Effect size” tab contains the important information you need to record.

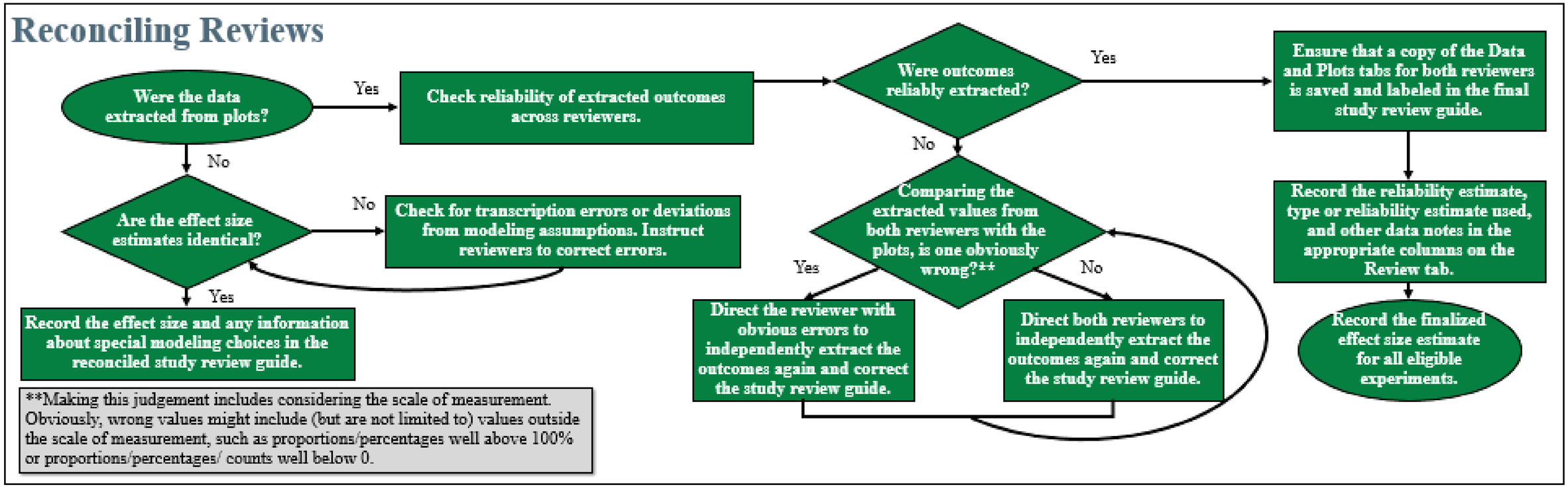
WWC-calculated findings

Design Comparable Effect Size Estimate	Change Sign of Effect for Meta-Analysis?	Standard Error	Autocorrelation Estimate	Note any deviations from default modeling guidance
9.99	No	1.22	-0.45	None

The design-comparable effect size infographic



The design-comparable effect size infographic



Questions?

Have questions? Contact us: https://ies.ed.gov/ncee/wwc/help

Mariola Moeyaert, Ph.D.
The State University of New York at Albany

Daniel M. Swan, Ph.D.
University of Oregon

Emily Tanner-Smith, Ph.D.
University of Oregon and
What Works Clearinghouse

References and further reading

- Dunlap, G., White, R., Vera, A., Wilson, D., & Panacek, L. (1996). The effects of multi-component, assessment-based curricular modifications on the classroom behavior of children with emotional and behavioral disorders. *Journal of Behavioral Education*, *6*(4), 481–500.
- 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. Retrieved October 2, 2020, from https://doi.org/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. Retrieved October 2, 2020, from https://doi.org/10.1002/jrsm.1086
- Laushey, K. M., & Heflin, L. J. (2000). Enhancing social skills of kindergarten children with autism through the training of multiple peers as tutors. *Journal of Autism Developmental Disorders, 30,* 183–193. Retrieved October 2, 2020, from https://doi.org/10.1023/A:1005558101038
- Moeyaert, M., Maggin, D., & Verkuilen, J. (2016). Reliability, validity, and usability of data extraction programs for single-case research designs. *Behavior Modification*, *40*(6), 874–900.
- Pustejovsky, J. E. (2016). *scdhlm: A web-based calculator for between-case standardized mean differences* (Version 0.3.1) [Web application]. Retrieved October 2, 2020, from https://jepusto.shinyapps.io/scdhlm
- 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. Retrieved October 2, 2020, from https://doi.org/10.3102/1076998614547577
- Saddler, B., Asaro-Saddler, K., Moeyaert, M., & Ellis-Robinson, T. (2017). Effects of a summarizing strategy on written summaries of children with emotional and behavioral disorders. *Remedial and Special Education*, *38(*2), 87–97. Retrieved October 2, 2020, from https://journals.sagepub.com/doi/10.1177/0741932516669051

References and further reading

- 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*(3–4), 528–553. Retrieved October 2, 2020, from https://www.tandfonline.com/doi/abs/10.1080/09602011.2013.819021
- Valentine, J. C., Tanner-Smith, E. E., & Pustejovsky, J. E. (2016). *Between-case standardized mean difference effect sizes for SCDs: A primer and tutorial using the scdhlm web application.* Oslo, Norway: The Campbell Collaboration. Retrieved October 2, 2020, from https://doi.org/10.4073/cmpn.2016.3
- What Works Clearinghouse, Institute of Education Sciences, U.S. Department of Education. (2020). *WWC version 4.1 standards and procedures handbooks.* Retrieved October 2, 2020, from https://ies.ed.gov/ncee/wwc/Handbooks