UNIVERSITY AT ALBANY
State University of New York

Usage of IPD Meta-Analysis for Synthesis of SCD Studies

Mariola Moeyaert
mmoeyaert@albany.edu
Associate Professor of Statistics
Educational Psychology and Methodology
University at Albany – SUNY
September 30, 2022

Dr. Mariola Moeyaert

UNIVERSITY AT ALBANY
State University of New York

The research reported in this presentation was supported by the Institute of Education Sciences, U.S. Department of Education, through grant R305D110024, R305D150007, and R305D190022.

The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education

- R305D110024, Multilevel synthesis of single-case experimental data: Further developments and empirical validation"
- R305D150007, Multilevel modeling of single-subject experimental data: Handling data and design complexities"
- **R305D190022, Assessing generalizability and variability of single-case design effect sizes using multilevel modeling including moderators**

Dr. Mariola Moeyaert

UNIVERSITY AT ALBANY
State University of New York

Acknowledgement

- Current research lab: Yukang Xue, Panpan Yan and Xinyun Xu
- Collaborators - IES grants: Wim Van den Noortgate, S. Natasha Beretvas, and John Ferron
- Collaborators – empirical demonstration: Karolina Urton, Richard Boon, Anna Barwasser, Kirsten Nobel and Matthias Grünke

Outline

UNIVERSITY AT ALBANY
State University of New York

1. Conceptual Framework
   1.1 Single-Case Research
   1.2 IPD Meta-Analysis
2. Demonstration: Usage of IPD Meta-Analysis using Real Data
3. Methodological work: Monte Carlo Simulation
4. Future research / Questions

1. Conceptual Framework
Single-case experimental designs

UNIVERSITY AT ALBANY
State University of New York

- SCDs are designed experiments in which one unit is observed repeatedly during a certain period of time under different levels of at least one manipulated variable.

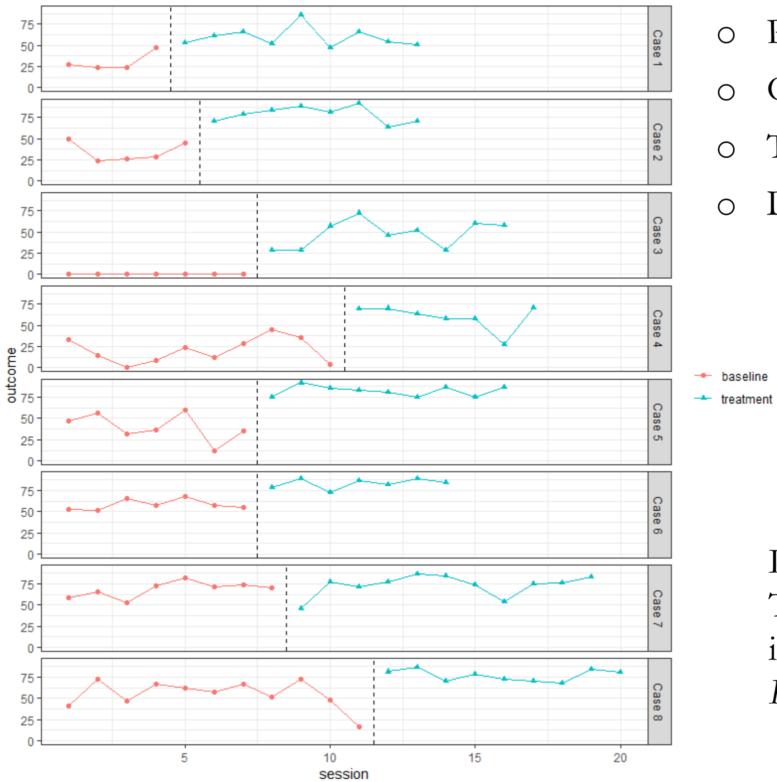
- SCDs are designs with the potential to demonstrate a causal effect.

- The basic SCD has many variations, but all SCDs often involve repeated, systematic measurement of a dependent variable before, during, and/or after the active manipulation of an independent variable.

1. Conceptual Framework
Single-case experimental designs

UNIVERSITY AT ALBANY
State University of New York

Example:



- Participants: 8 autistic children (4 nonverbal and 4 echolalic).
- Outcome: Frequency of verbalizations
- Treatment: Natural Language Paradigm (NLP)
- Design: Multiple Baseline Design (MBD) across participants

Laski, K. E., Charlop, M. H., & Schreibman, L. (1988). Training parents to use the natural language paradigm to increase their autistic children's speech. *Journal of Applied Behavior Analysis*, *21*, 391-400.

1. Conceptual Framework

Single-case experimental designs

<u>Used in a variety of different fields</u>

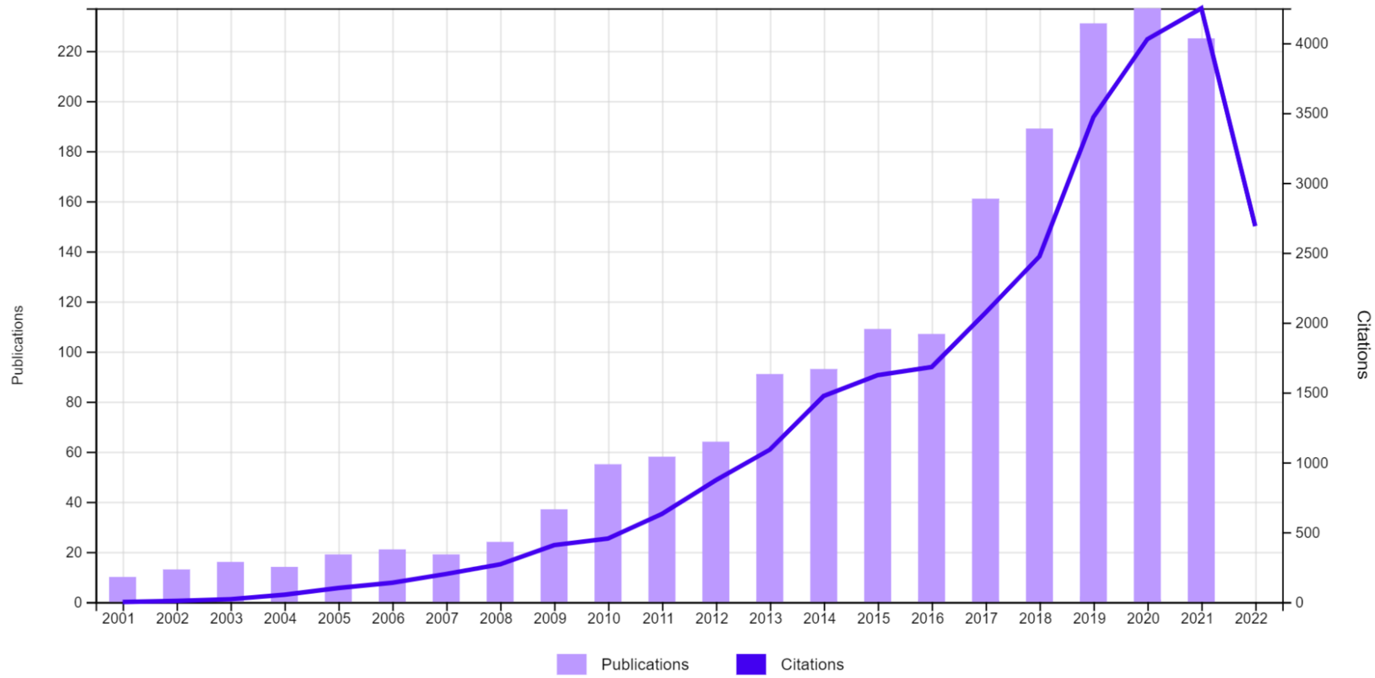
Source: Web Of Science, Sept 2022, Keywords:
TS=(single-case* OR single-subject* OR interrupted time series* of intra-subject* or n-of-1*)

1. Conceptual Framework

Single-case experimental designs

UNIVERSITY AT ALBANY
State University of New York

Exponential increase in popularity



Source: Web Of Science, Keywords: TS=(single-case* OR single-subject* OR interrupted time series* of intra-subject* or n-of-1*), field = Education.

1. Conceptual Framework
Single-case experimental designs

UNIVERSITY AT ALBANY
State University of New York

Evidence-Based Education Policy

- We have entered an era in which scientific evidence will increasingly inform policy.
- Combining evidence from multiple SCD studies, using **meta-analytic techniques**, can provide a basis for generalization about effects of intervention.

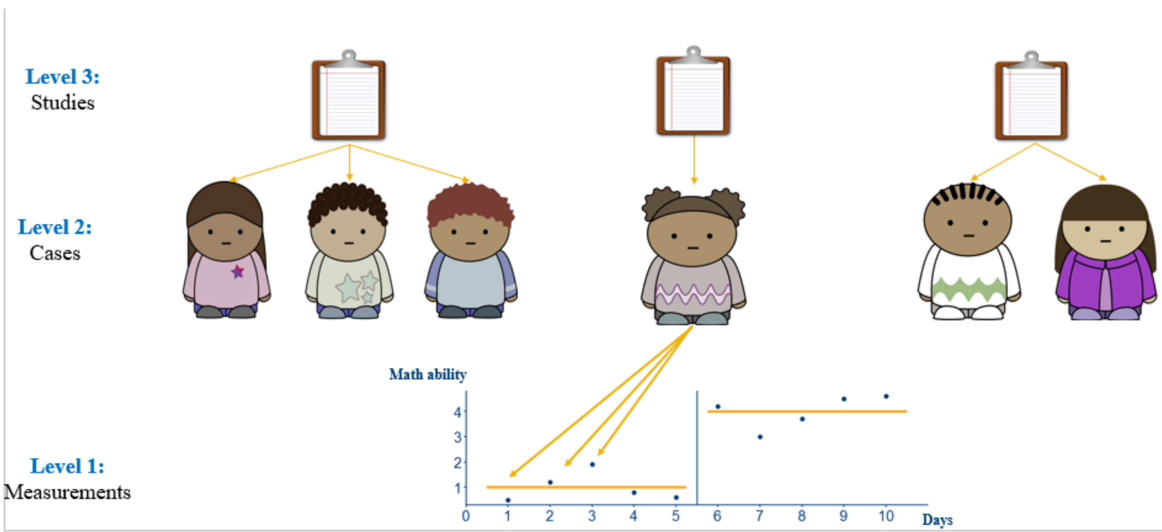
- Using **meta-analysis**, the focus is on
  - ***Summarizing magnitude*** of intervention effects.
  - Investigating ***intervention heterogeneity***
  - Identifying ***moderators*** to explain intervention heterogeneity

1. Conceptual Framework
Single-case experimental designs

UNIVERSITY AT ALBANY
State University of New York

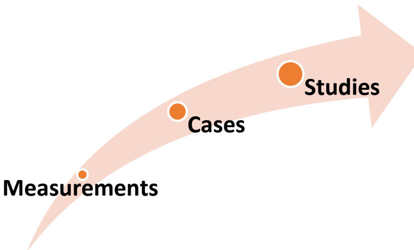
IPD Meta-Analysis

- Raw SCD data meta-analysis is also called raw Individual Patient/Participant Data (IPD) meta-analysis (Declercq et al. 2020, Moeyaert & Fingerhut, 2022).
- Raw data from multiple participants and studies are synthesized.
- Three-level structure:



1. Conceptual Framework

IPD Meta-Analysis



<u>Why IPD Meta-Analysis?</u>

- Takes the hierarchical nature of the data into account.
- Estimate of the overall intervention effect across cases

and across studies in addition to participant-specific and study-specific treatment effects.

- Between-participant and between-study variance can be estimated.
- Flexibility: investigating moderators at the case and study level to explain intervention heterogeneity.

Moeyaert, M., & Yang, P. (2021). Assessing generalizability and variability of single-case design effect sizes using two-stage multilevel modeling including moderators, *Behaviormetrika*, *48*, 207-229. Doi: 10.1007/s41237-021-00141-z

1. Conceptual Framework

IPD Meta-Analysis

UNIVERSITY AT ALBANY
State University of New York

Statistical Model - Two-stage IPD meta-analysis

[Declercq et al., 2020; Moeyaert & Fingerhut, 2022; Moeyaert & Yang, 2021]

- Stage one: pre-processing step

In the first stage, effect size(s) need to be estimated from the raw IPD data.

Pre-Processing model: $y_{ijk} = \beta_{0jk} + \beta_{1jk} Trt_{ijk} + e_{ijk}$ with $e_{ijk} \sim N\,(0, \sigma_e^2)$

  - This provides an estimate of the participant-specific regression coefficient reflecting the effect size, b_{1jk}, and the within-participant residual variance, σ_e^2.

- Stage two: IPD meta-analysis
  - The effect size b_{1jk} can be used in the three-level meta-analytical model.

1. Conceptual Framework
IPD Meta-Analysis

UNIVERSITY AT ALBANY
State University of New York

Statistical Model - IPD Meta-Analysis approaches

One-stage IPD meta-analysis

$$y_{ijk} = \beta_{0jk} + \beta_{1jk}Trt_{ijk} + e_{ijk}$$

$$\begin{cases} \beta_{0jk} = \boxed{\gamma_{000}} + u_{0jk} + v_{00k} \\ \beta_{1jk} = \boxed{\gamma_{100}} + u_{1jk} + v_{10k} \end{cases}$$

$$\begin{pmatrix} u_{0jk} \\ u_{1jk} \end{pmatrix} \sim MVN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \boxed{\begin{pmatrix} \sigma^2_{u_0} & \\ \sigma_{u_{01}} & \sigma^2_{u_1} \end{pmatrix}}\right]$$

$$\begin{pmatrix} v_{0jk} \\ v_{1jk} \end{pmatrix} \sim MVN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \boxed{\begin{pmatrix} \sigma^2_{v_0} & \\ \sigma_{v_{01}} & \sigma^2_{v_1} \end{pmatrix}}\right]$$

Two-stage IPD meta-analysis

$$y_{ijk} = \beta_{0jk} + \beta_{1jk}Trt_{ijk} + e_{ijk}$$

$$b_{1jk} = \beta_{1jk} + r_{jk}$$

$$\beta_{1jk} = \boxed{\gamma_{100}} + u_{1jk} + v_{10k}$$

$$u_{1jk} \sim N\left(0, \boxed{\sigma^2_{u_{1jk}}}\right)$$

$$v_{10k} \sim N\left(0, \boxed{\sigma^2_{v_{10k}}}\right)$$

→One stage approach: two fixed effects and seven random effect estimates

→Two-stage approach: one fixed effect and two random effect estimates

1. Conceptual Framework
IPD Meta-Analysis

UNIVERSITY AT ALBANY
State University of New York

Statistical Model - IPD Meta-Analysis approaches

- For more complex models, the three-level models involve more regression coefficients and therefore more parameters to estimate. This is particularly true for the variance components, since the dimensions of the covariance matrices at the higher level(s) increase quickly.

- The two-stage approach has an important potential benefit over the one-stage approach when the underlying model is more complex. The multilevel model estimated based on the effect sizes is reduced, so there are **less parameters to estimate.** This results in faster estimation procedures and better convergence rates compared to the one-stage approach.

- The precision and the bias of the point estimates is very similar for both approaches (Declercq et al., 2020)

1. Conceptual Framework

IPD Meta-Analysis

UNIVERSITY AT ALBANY
State University of New York

Statistical Model - Two-stage IPD meta-analysis

Pre-Processing model: $Y_{ijk} = \beta_{0jk} + \beta_{1jk}Trt_{ijk} + e_{ijk}\ with\ e_{ijk} \sim N(0, \sigma_e^2)$

- b_{1jk} is a function of the true participant-specific effect size β_{1jk} and the residual standard deviation is assumed to be known (obtained from the pre-processing step):

Level 1 – Observation Level: $b_{1jk} = \beta_{1jk} + r_{1jk}$

1. Conceptual Framework

IPD Meta-Analysis

UNIVERSITY AT ALBANY
State University of New York

Statistical Model - Two-stage IPD meta-analysis

Level 2: **Variation between participants from the same study**

$$\beta_{1jk} = \theta_{10k} + u_{1jk} \text{ with } u_{1jk} \sim N\left(0, \sigma^2_{u_{1jk}}\right)$$

Level 3: **Variation between studies**

$$\theta_{10k} = \gamma_{100} + v_{10k} \text{ with } v_{10k} \sim N\left(0, \sigma^2_{v_{10k}}\right)$$

- Meta-analysts are interested in the **estimate of $\boldsymbol{\gamma_{100}}$**, which expresses the overall intervention effect across participants and across studies, and in the variance component $\boldsymbol{\sigma^2_{u_1}}$, which expresses the extent to which the intervention effect varies across participants within the study, and the variance component $\boldsymbol{\sigma^2_{v_1}}$, expressing the extend to which the intervention varies across studies.

1. Conceptual Framework

IPD Meta-Analysis

UNIVERSITY AT ALBANY
State University of New York

Statistical Model

Combined Model:

Overall intervention Effect

$$b_{1jk} = \gamma_{100} + v_{10k} + u_{1jk} + r_{1jk}$$

Between-case variance in intervention effect

Between-study variance in intervention effect

with $r_{ijk} \sim N\,(0, \sigma_r^2)$, and $u_{1jk} \sim N\left(0, \sigma_{u_{1jk}}^2\right)$ and $v_{10k} \sim N\left(0, \sigma_{v_{10k}}^2\right)$

1. Conceptual Framework

IPD Meta-Analysis: Standardization

UNIVERSITY AT ALBANY
State University of New York

- Example of two studies included in Urton et al. (in preparation)
  - Onachukwu et al. (2007), DV = percentage (percent correct on the reading comprehension tests)
  - Calvin et al. (2022), DV = composite score from rate and accuracy subtests
- Therefore, standardization is needed (Van den Noortgate & Onghena, 2008, Moeyaert et al., 2013).
- The participant-specific effect sizes ($b'_{1jk}s$)are standardized by dividing them by the estimated residual within-subject standard deviation of participant j from study k, $\hat{\sigma}_{ejk}$ (Van den Noortgate & Onghena, 2008):

$$b'_{1jk} = \frac{b_{1jk}}{\hat{\sigma}_{ejk}}$$

$\hat{\sigma}_{ejk}$ is obtained by running the following OLS per participant :

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk}Trt_{ijk} + e_{ijk}\ with\ e_{ijk} \sim N(0, \sigma_e^2)$$

[This is already part of the preprocessing step]

1. Conceptual Framework

IPD Meta-Analysis: Bias Correction

UNIVERSITY AT ALBANY
State University of New York

- Hedges's correction procedure is promising for reducing the bias of the fixed effect estimation for combining the standardized regression coefficients, b'_{1jk} (Ugille et al., 2014, Jamshidi et al., 2021).

$$\left(b'_{1jk}\right)^c = b'_{1jk}\left(1 - \frac{3}{4m-1}\right)$$

With m = df = [I – p – 1]

When the bias-correction factor is applied to the standardized regression coefficient estimates, it should also be applied to the standard error estimates (and therefore the sampling variance) associated with each of the coefficients:

$$\left((\hat{\sigma}_{b'})^c\right)^2 = \widehat{\sigma^2_{b'}}\left(1 - \frac{3}{4m-1}\right)^2$$

- Next, the bias-corrected standardized effect sizes can be synthesized.

Jamshidi, L., *Declercq, L., Fernández-Castilla, B., Ferron, J., **Moeyaert, M.**, Beretvas, S.N., & Van den Noortgate, W. (2021). Bias adjustment in multilevel meta-analysis of standardized single-case experimental data. *Journal of Experimental Education*, 89, 334-361. https://doi.org/10.1080/00220973.2019.1658568

1. Conceptual Framework

IPD Meta-Analysis: Moderators

UNIVERSITY AT ALBANY
State University of New York

Goal of using two-stage IPD meta-analysis

Can provide an answer to a variety of interesting questions:

- What is the overall average treatment effect? [$\hat{\gamma}_{100}$]
- Does the size of the intervention effect vary across participants? [$\hat{\sigma}^2_{u_1}$]
- Does the size of the intervention effects (change in level) vary across studies? [$\hat{\sigma}^2_{v_2}$]
- What participant factors relate to effect size? [moderators at level 2 can be added]?
- What study factors relate to effect size? [moderators at level 3 can be added]?

1. Conceptual Framework

IPD Meta-Analysis: Moderators

UNIVERSITY AT ALBANY
State University of New York

Intervention heterogeneity

In order to explain intervention heterogeneity between cases and/or studies, moderators at levels 2 and/or level 3 of the IPD meta-analytic model can be added

Level 1: $\left(b'_{1jk}\right)^{c} = \beta_{1jk} + r_{1jk}$ with $r_{1jk} \sim N\,(0, \sigma^2_{r1jk})$

Level 2: $\beta_{1jk} = \theta_{10k} + \sum_{p=1}^{P} \theta_{1pk}\, Z_{1pk} + u_{1jk}$ with $u_{2jk} \sim N\left(0, \sigma^2_{u_{1jk}}\right)$

Level 3: $\theta_{10k} = \gamma_{100} + \sum_{q=1}^{Q} \gamma_{10q}\, B_{10q} + v_{10k}$ with $v_{10k} \sim N\left(0, \sigma^2_{v_{10k}}\right)$

* with *P* referring to the number of predictors at the second level and Q referring the number of third level predictors.

2. Demonstration

Urton, K. Boon, R. Moeyaert, M., Barwasser, A., Nobel, K., & Grünke, M *(in preparation). Effects of Graphic Organizer Interventions on Competencies for At-Risk Students and Students with Disabilities: A Three-Level Meta-Analysis of Single-Case Data*

- Outcome: Competences such as reading comprehension, writing, and listening comprehension
- Intervention: Graphic organizers such as concept maps, cognitive maps, semantic maps, story maps and Venn diagrams as well as schematic representations.
- 40 primary SCD studies, and 159 participants

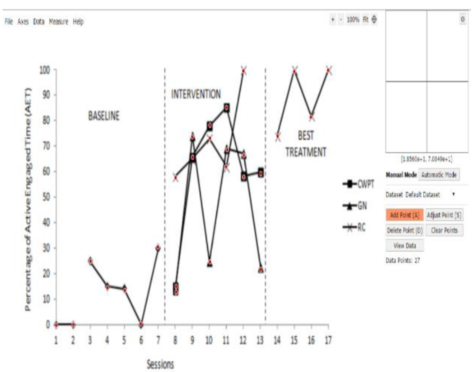
[Inclusion criteria: (a) used a graphic organizer as the primary intervention either alone, paired with another strategy or as part of an instructional package, (b) included at least three participants identified with a disability, (c) took place in a K-12 classroom in a school in the United States, (d) employed a single-case research design, (e) were written in English, and (f) were published in peer-reviewed journals between 1975 and July 25, 2022.]

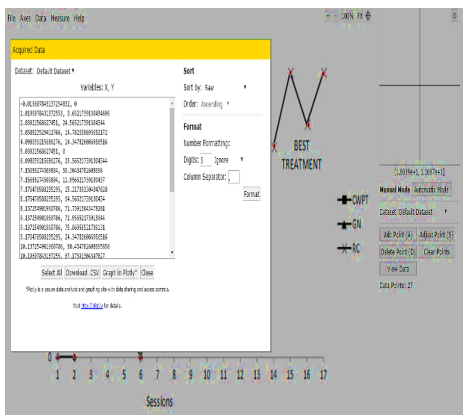
2. Demonstration

UNIVERSITY AT ALBANY
State University of New York

- Raw SCD data from all participant graphs were retrieved.
- The data retrieval software program, WebPlotDigitizer 4.6 (Rohatgi, 2022), was used for this purpose.

Participant graph → Raw data retrieval → Dataset





Study	Case	Phase	Time	Y
Bambara1995	Al	0	1	3
Bambara1995	Al	0	2	3
Bambara1995	Al	0	3	2
Bambara1995	Al	0	4	4
Bambara1995	Al	1	5	1
Bambara1995	Al	1	6	1
Bambara1995	Al	1	7	0
Bambara1995	Al	1	8	0

2. Demonstration

UNIVERSITY AT ALBANY
State University of New York

	A	B	C	D	E	F	G	I	J	N	O	P	S	T	U	V	W	X	Y
1	Study_Name	Study	Case	Session	Outcomes	Time	D	Design	Type_Outcome	Interventionist	Digital	isability_1	Quality	_integrity	Dosage	Social_Validity	Age	Grade	Gender
2	1.Alves et al., 2015	1	1	0.937	41.822	1	0	1	1	1	0	1	1	1	10	0	10	5	0
3	1.Alves et al., 2015	1	1	1.937	41.822	2	0	1	1	1	0	1	1	1	10	0	10	5	0
4	1.Alves et al., 2015	1	1	2.988	-0.428	3	0	1	1	1	0	1	1	1	10	0	10	5	0
5	1.Alves et al., 2015	1	1	3.887	83.555	4	1	1	1	1	0	1	1	1	10	0	10	5	0
6	1.Alves et al., 2015	1	1	4.899	83.208	5	1	1	1	1	0	1	1	1	10	0	10	5	0
7	1.Alves et al., 2015	1	1	5.912	83.033	6	1	1	1	1	0	1	1	1	10	0	10	5	0
8	1.Alves et al., 2015	1	1	6.851	93.389	7	1	1	1	1	0	1	1	1	10	0	10	5	0
9	1.Alves et al., 2015	1	1	7.815	103.395	8	1	1	1	1	0	1	1	1	10	0	10	5	0
10	1.Alves et al., 2015	1	1	8.839	93.046	9	1	1	1	1	0	1	1	1	10	0	10	5	0
11	1.Alves et al., 2015	1	1	9.864	92.697	10	1	1	1	1	0	1	1	1	10	0	10	5	0
12	1.Alves et al., 2015	1	1	10.839	103.046	11	1	1	1	1	0	1	1	1	10	0	10	5	0
13	1.Alves et al., 2015	1	1	11.840	92.701	12	1	1	1	1	0	1	1	1	10	0	10	5	0
14	1.Alves et al., 2015	1	1	12.816	92.360	13	1	1	1	1	0	1	1	1	10	0	10	5	0
15	1.Alves et al., 2015	1	1	14.879	49.763	14	3	1	1	1	0	1	1	1	10	0	10	5	0
16	1.Alves et al., 2015	1	2	0.934	52.692	1	0	1	1	1	0	5	1	1	10	0	11	5	1
17	1.Alves et al., 2015	1	2	1.948	31.731	2	0	1	1	1	0	5	1	1	10	0	11	5	1
18	1.Alves et al., 2015	1	2	2.910	42.308	3	0	1	1	1	0	5	1	1	10	0	11	5	1
19	1.Alves et al., 2015	1	2	3.857	84.615	4	1	1	1	1	0	5	1	1	10	0	11	5	1
20	1.Alves et al., 2015	1	2	4.859	52.692	5	1	1	1	1	0	5	1	1	10	0	11	5	1
21	1.Alves et al., 2015	1	2	5.807	84.615	6	1	1	1	1	0	5	1	1	10	0	11	5	1
22	1.Alves et al., 2015	1	2	6.808	63.269	7	1	1	1	1	0	5	1	1	10	0	11	5	1
23	1.Alves et al., 2015	1	2	8.758	73.846	8	1	1	1	1	0	5	1	1	10	0	11	5	1
24	1.Alves et al., 2015	1	2	9.744	84.615	9	1	1	1	1	0	5	1	1	10	0	11	5	1
25	1.Alves et al., 2015	1	2	10.706	95.385	10	1	1	1	1	0	5	1	1	10	0	11	5	1
26	1.Alves et al., 2015	1	2	11.668	105.385	11	1	1	1	1	0	5	1	1	10	0	11	5	1
27	1.Alves et al., 2015	1	2	12.669	95.192	12	1	1	1	1	0	5	1	1	10	0	11	5	1
28	1.Alves et al., 2015	1	2	14.631	105.385	13	3	1	1	1	0	5	1	1	10	0	11	5	1
29	1.Alves et al., 2015	1	3	1.000	70.702	1	0	1	1	1	0	9	1	1	10	0	9	3	0
30	1.Alves et al., 2015	1	3	2.000	70.702	2	0	1	1	1	0	9	1	1	10	0	9	3	0
31	1.Alves et al., 2015	1	3	3.000	20.175	3	0	1	1	1	0	9	1	1	10	0	9	3	0
32	1.Alves et al., 2015	1	3	4.015	80.526	4	0	1	1	1	0	9	1	1	10	0	9	3	0
33	1.Alves et al., 2015	1	3	5.000	40.175	5	0	1	1	1	0	9	1	1	10	0	9	3	0
34	1.Alves et al., 2015	1	3	5.925	80.702	6	0	1	1	1	0	9	1	1	10	0	9	3	0

2. Demonstration

UNIVERSITY AT ALBANY
State University of New York

<u>IPD Meta-Analytic Model</u>:

Effect of self-monitoring interventions on writing outcomes

$$\left(b'_{1jk}\right)^{c} = \gamma_{100} + v_{10k} + u_{1jk} + r_{1jk}$$

with $r_{ijk} \sim N\left(0, \sigma_r^2\right)$, and $u_{1jk} \sim N\left(0, \sigma_{u_{1jk}}^2\right)$ and $v_{10k} \sim N\left(0, \sigma_{v_{10k}}^2\right)$

<u>Results</u>

- There is a statistically significant increase in writing outcomes after exposure to the intervention [$\hat{\gamma}_{100} = 3.419, SE = 0.474, t(37.9) = 7.22, p < .0001$]. There is an increase in performance of 3.36 standardized units.
- There is some evidence for variability in intervention effectiveness (i.e., intervention heterogeneity) between cases ($\hat{\sigma}_{u_1}^2 = 4.699, SE = 0.670, Z = 7.02, p < .0001$), and between studies ($\hat{\sigma}_{v_1}^2 = 7.583, SE = 2.078, Z = 3.65, p = .0001$).

2. Demonstration

UNIVERSITY AT ALBANY
State University of New York

IPD Meta-Analytic Model:

Explaining intervention heterogeneity by including participant and study level moderators.

- Study-level moderators

Quality

Quality_r	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	16	40.00	16	40.00
1	24	60.00	40	100.00

Integrity

Treatment_integrity	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	7	17.50	7	17.50
1	33	82.50	40	100.00

Digital

Digital	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	29	72.50	29	72.50
1	11	27.50	40	100.00

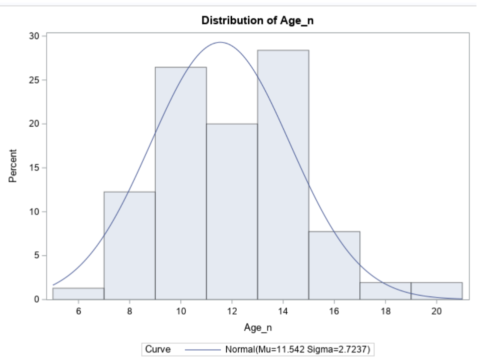
2. Demonstration

UNIVERSITY AT ALBANY
State University of New York

IPD Meta-Analytic Model:

- Participant-level moderators

Age (# missing = 4)



Basic Statistical Measures			
Location		Variability	
Mean	11.54194	Std Deviation	2.72373
Median	11.00000	Variance	7.41868
Mode	13.00000	Range	15.00000
		Interquartile Range	3.00000

Disability incidence

incidence	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	10	6.29	10	6.29
1	149	93.71	159	100.00

Other moderators have many missing values:

- Dosage (# missing = 80)
- IQ (# missing = 75)

2. Demonstration

UNIVERSITY AT ALBANY
State University of New York

<u>IPD Meta-Analytic Model</u>:

Explaining intervention heterogeneity by including participant and study level moderators.

$$\left(b'_{1jk}\right)^{c} = \gamma_{100}$$
$$+ \gamma_{101}Age\_c_{101} + \gamma_{101}Incidence_{102}$$
$$+ \gamma_{110}Digital_{11k} + \gamma_{120}Quality_{12k}$$
$$+ v_{10k} + u_{1jk} + \ r_{1jk}$$

*Continuous variable Age was mean-centered (around 11.54) prior to analysis to provide meaningful interpretations.

2. Demonstration

UNIVERSITY AT ALBANY
State University of New York

IPD Meta-Analytic Model:

Explaining intervention heterogeneity by including participant and study level moderators.

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
D_total	2.5611	1.7847	65	1.44	0.1561
Age_c	0.2125	0.1481	85.7	1.43	0.1550
incidence	0.4210	1.7198	78.9	0.24	0.8072
Digital	0.4460	1.1734	35.6	0.38	0.7062
Quality_r	0.5281	1.0297	34.4	0.51	0.6113

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
D_total	2.9407	0.8788	35.3	3.35	0.0020
Age_c	0.2146	0.1469	84.8	1.46	0.1477
Digital	0.4588	1.1596	36	0.40	0.6947
Quality_r	0.5471	1.0158	34.8	0.54	0.5936

Can we "trust" these results? Are the intervention and moderator effect estimates unbiased/precise? Are the standard errors unbiased? Is there sufficient power to estimate true intervention and moderator effects?

3. Methodological Research

Mote Carlo Simulation Study

UNIVERSITY AT ALBANY
State University of New York

Moeyaert, M., Yang, P., & Xue, Y. (in preparation). Three-level meta-analysis of single-case research including moderators: Empirical validation using a large-scale Monte-Carlo simulation study.

Purpose:

- Methodological work is needed to empirically investigate under which realistic SCED conditions (e.g., number of measurement occasions, participants, magnitude of the effects and variance) intervention and moderator effects can be estimated with appropriate statistical properties.

3. Methodological Research

Simulation Study – Data Generation

UNIVERSITY AT ALBANY
State University of New York

- SCED studies commonly include 0 to 2 moderators at levels two and three (Moeyaert et al., 2022). Most commonly used measurement scale is nominal (with two categories), and most commonly used combination is two nominal variables and one continuous variable.

Moeyaert, M., Yang, P., Xu, X., & Kim, E. (2021). Characteristics of moderators in meta-analyses of single-case experimental design studies. *Behavior Modification*. https://doi.org/10.1177/01454455211002111

3. Methodological Research

Simulation Study – Data Generation

UNIVERSITY AT ALBANY
State University of New York

Table 2. Overview of Moderator Characteristics Reported by at Least Five Meta-Analytic Studies.

Moderator level	Moderator	Number (%)[a]	Measurement scale	Missing data[b]	Analysis approach (%)[c]
Study level	Study design	42 (70)	Nominal	No ($n=29$); yes ($n=2$); not clear ($n=11$)	*D* (95); *F* (90); *Q* (52); *S* (29)
	Physical setting of intervention	38 (63)	Nominal	No ($n=28$); yes ($n=1$); not clear ($n=9$)	*D* (97); *F* (87); *Q* (45); *S* (37)
	Design standards	26 (43)	Nominal	No ($n=22$); yes ($n=3$); not clear ($n=1$)	*D* (100); *F* (85); *Q* (27); *S* (12)
	Design strength	23 (38)	Nominal	No ($n=17$); yes ($n=1$); not clear ($n=5$)	*D* (87); *F* (83); *Q* (26); *S* (22)
	Interobserver agreement	15 (25)	Nominal ($n=11$); continuous ($n=4$)	No ($n=7$); yes ($n=5$); not clear ($n=3$)	*D* (100); *F* (60); *Q* (7)
	Maintenance	15 (25)	Nominal	No ($n=2$); yes ($n=10$); not clear ($n=3$)	*D* (100); *F* (80); *Q* (33); *S* (13)
	Generalization	13 (22)	Nominal	No ($n=2$); yes ($n=11$)	*D* (100); *F* (85); *Q* (31)
	Instructional arrangement	11 (18)	Nominal	No ($n=7$); yes ($n=1$); not clear ($n=3$)	*D* (100); *F* (100); *Q* (64); *S* (45)
	Publication type	11 (18)	Nominal	No ($n=7$); not clear ($n=4$)	*D* (91); *F* (82); *Q* (55); *S* (36)
	Social/internal validity	10 (17)	Nominal ($n=9$); continuous ($n=1$)	No ($n=3$); yes ($n=5$); not clear ($n=2$)	*D* (100); *F* (60); *Q* (20)
	Context	6 (10)	Nominal	No ($n=3$); yes ($n=1$); not clear ($n=2$)	*D* (100); *F* (83); *Q* (50); *S* (33)
	FBA method	6 (10)	Nominal	No ($n=4$); yes ($n=1$); not clear ($n=1$)	*D* (100); *F* 83); *Q* (66); *S* (17)
	Improvement/ findings	5 (8)	Nominal	No ($n=3$); yes ($n=2$)	*D* (100); *F* (100); *Q* (40)
Participant level	Age	55 (92)	Nominal ($n=25$); ordinal ($n=8$); continuous ($n=25$)	No ($n=30$); yes ($n=12$); not clear ($n=13$)	*D* (100); *F* (87); *Q* (67); *S* (49)
	Disability status	44 (73)	Nominal	No ($n=22$); yes ($n=9$); not clear ($n=13$)	*D* (98); *F* (89); *Q* (57); *S* (45)
	Gender	39 (65)	Nominal	No ($n=22$); yes ($n=11$); not clear ($n=6$)	*D* (100); *F* (87); *Q* (33); *S* (21)
	Ethnicity	17 (26)	Nominal	No ($n=2$); yes ($n=13$); not clear ($n=2$)	*D* (100); *F* (71); *Q* (18); *S* (6)
	Functional repertories	17 (26)	Nominal ($n=14$); continuous ($n=3$)	No ($n=6$); yes ($n=8$); not clear ($n=3$)	*D* (100); *F* (76); *Q* (53); *S* (29)
	Received special education	5 (8)	Nominal	No ($n=2$); yes ($n=1$); not clear ($n=2$)	*D* (80); *F* (40); *Q* (60); *S* (60)

(continued)

3. Methodological Research

UNIVERSITY AT ALBANY
State University of New York

Simulation Study – Data Generation

Table 2. (continued)

Moderator level	Moderator	Number (%)[a]	Measurement scale	Missing data[b]	Analysis approach (%)[c]
Within-participant level	Intervention program[d]	44 (73)	Nominal	No ($n=26$); yes ($n=7$); not clear ($n=11$)	*D* (98); *F* (91); *Q* (73); *S* (50)
	Intervention agent[d]	28 (47)	Nominal	No ($n=13$); yes ($n=8$); not clear ($n=7$)	*D* (96); *F* (100); *Q* (57); *S* (43)
	Intervention techniques[d]	21 (35)	Nominal	No ($n=13$); yes ($n=3$); not clear ($n=5$)	*D* (100); *F* (100); *Q* (62); *S* (43)
	Intervention dosage[d]	18 (30)	Nominal ($n=10$); ordinal ($n=2$); continuous ($n=6$)	No ($n=5$); yes ($n=4$); not clear ($n=9$)	*D* (94); *F* (89); *Q* (50); *S* (39)
	Fidelity[d]	14 (23)	Nominal ($n=13$); continuous ($n=1$)	No ($n=3$); yes ($n=7$); not clear ($n=4$)	*D* (93); *F* (64); *Q* (14); *S* (14)
	Technology devices[d]	7 (12)	Nominal	No ($n=4$); yes ($n=1$); not clear ($n=2$)	*D* (86); *F* (71); *Q* (43); *S* (14)
	Outcome domain[e]	54 (90)	Nominal	No ($n=35$); yes ($n=5$); not clear ($n=14$)	*D* (100); *F* (93); *Q* (78); *S* (59)
	Methods of measuring outcomes[e]	14 (23)	Nominal	No ($n=9$); yes ($n=1$); not clear ($n=4$)	*D* (100); *F* (93); *Q* (36); *S* (21)

Note. *D* = description in words; *F* = frequency or frequency table; *Q* = quantitative metric; *S* = statistical analysis (reporting significance).
[a]Percent of SCED meta-analytic studies including this moderator.
[b]Number of SCED meta-analytic studies reporting degree of missing data for each moderator: no = reporting no missing data for one moderator; yes = reporting having missing data for one moderator; not clear = did not mentioning the information of missing data for one moderator.
[c]Percent of SCED meta-analytic studies using different types of analysis approach.
[d]Intervention specific moderators.
[e]Outcome specific moderators.

3. Methodological Research

Simulation Study – Data Generation

- Four models are used to generate MBD data, based on the following combined model:

$$y_{ijk} = \gamma_{000} + v_{00k} + u_{0jk} + \left(\gamma_{100} + \sum_{p=1}^{P} \gamma_{1p0} Z_{1pk} + \sum_{q=1}^{Q} \gamma_{10q} W_{10q} + v_{10k} + u_{1jk}\right) Trt_{ijk} + e_{ijk}$$

with $\begin{bmatrix} v_{00k} \\ v_{10k} \end{bmatrix} \sim MVN(0, \Sigma_v)$, $\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim MVN(0, \Sigma_u)$, and $e_{ij} \sim N(0, \sigma_e^2)$

3. Methodological Research

Simulation Study – Data Generation

UNIVERSITY AT ALBANY
State University of New York

- Model 0: No moderators
- Model 1: 1 level-2 (Gender) and 1 level-3 (Quality) moderator
- Model 2: 2 level-2 (Gender and Age) and 1 level-3 (Quality) moderators
- Model 3: 2 level-2 (Gender and Age) and 2 level-3 (Quality and Setting) moderators

3. Methodological Research

Simulation Study – Data Generation

UNIVERSITY AT ALBANY
State University of New York

Table

Design conditions and realistic parameters

Design Factor	Notation	Value
Number of studies	K	10, 30, 40 or 50
Number of observations	I	20 or 40
Number of participants	J	4, 7, 12 or 20
Intervention effect	γ_{100}	0 or 2.00
Moderator effects	Gender, γ_{110}	0.75, 1.00, 1.50, 2.00
	Age, γ_{120}	0.25 or 0.50
	Study Quality, γ_{101}	0.75, 1.00, 1.50, 2.00
	Physical setting, γ_{102}	0.75, 1.00, 1.50, 2.00
Between-case variance	Baseline level, $\sigma_{\theta 0}^2$	2.00
	Intervention effect, $\sigma_{\theta 1}^2$	2.00
Between-study variance	Baseline level, $\sigma_{\theta 0}^2$	2.00
	Intervention effect, $\sigma_{\theta 1}^2$	2.00
Within-participant variance	σ_e^2	1.00

3. Methodological Research

Simulation Study – Data Generation

UNIVERSITY AT ALBANY
State University of New York

- The number of conditions investigated depends on the specific model of interest.
- Model 0 is the only model that does not include 40 or 50 studies. The reason for this is that statistical properties are appropriate with as few as 30 studies (and there is already sufficient power across all conditions with 30 studies).
- Number of conditions per model:
  - Model 0: 2 x 2 x 4 x 2 = 32 conditions;
  - Model 1: 4 x 2 x 4 x 2 x 4 x 4 = 1,024 conditions;
  - Model 2: 4 x 2 x 4 x 2 x 4 x 2 x 4 = 2,048 conditions;
  - Model 3: 4 x 2 x 4 x 2 x 4 x 2 x 4 x 4 = 8,192 conditions
- For each condition, 1,000 datasets are examined. This resulted in a total of 11,296,000 datasets to be examined [(32 + 1,024 + 2,048 + 8,192) x 1,000 = 11,296,000].

3. Methodological Research

Simulation Study – Data Estimation

UNIVERSITY AT ALBANY
State University of New York

- All models, pre-Processing model:

$$y_{ijk} = \beta_{0jk} + \beta_{1jk}Trt_{ijk} + e_{ijk} \text{with } e_{ijk} \sim N\,(0, \sigma_e^2)$$

$$b'_{1jk} = \frac{b_{1jk}}{\hat{\sigma}_{ejk}}$$

$$\left(b'_{1jk}\right)^C = b'_{1jk}\left(1 - \frac{3}{4m-1}\right) \text{ and } \left((\hat{\sigma_{b'}})^c\right)^2 = \widehat{\sigma^2_{b'}}\left(1 - \frac{3}{4m-1}\right)^2$$

- Model 0: $\left(b'_{1jk}\right)^c = \gamma_{100} + v_{10k} + u_{1jk} + r_{1jk}$
- Model 1: $\left(b'_{1jk}\right)^c = \gamma_{100} + \gamma_{101}\, Quality_{101} + \gamma_{110}Gender_{11k} + v_{10k} + u_{1jk} + r_{1jk}$
- Model 2: $\left(b'_{1jk}\right)^c =$
  $\gamma_{100} + \gamma_{101}\, Quality_{101} + \gamma_{110}Gender_{11k} + \gamma_{120}Age_{12k} + v_{10k} + u_{1jk} + r_{1jk}$
- Model 3: $\left(b'_{1jk}\right)^c =$
  $\gamma_{100} + \gamma_{101}\, Quality_{101} + \gamma_{101}Setting_{102} + \gamma_{110}Gender_{11k} + \gamma_{120}Age_{12k} + v_{10k} + u_{1jk} +$
  r_{1jk}

3. Methodological Research

Simulation Study - Analysis

UNIVERSITY AT ALBANY
State University of New York

- Generalized linear modeling (GLM) is used to identify design factors that have a statistically significant and large (η_p^2) impact on the following statistical properties.

$Bias\ (\hat{\gamma}_{100}) = \bar{\hat{\gamma}}_{100} - \gamma_{100}$

$Rel\ Bias\ (\hat{\gamma}_{100}) = (\bar{\hat{\gamma}}_{100} - \gamma_{100})/\gamma_{100}$

MSE $(\hat{\gamma}_{100}) = [Bias\ (\hat{\gamma}_{100})]^2$ + Variance

$Rel\ Standard\ Error\ Bias\ (\hat{\gamma}_{200})$ =
$(Median\ of\ SE\ of\ \hat{\gamma}_{200}\ STD\ of\ distribution\ of\ \hat{\gamma}_{200})/STD\ of\ distribution\ of\ \hat{\gamma}_{200}$

$$SE\_Bias\ (\hat{S}_{\hat{\gamma}_{100}}) = \frac{\bar{\hat{S}}_{\hat{\gamma}_{100}} - \hat{S}_{\hat{\gamma}_{100\_EMP}}}{\hat{S}_{\hat{\gamma}_{100\_EMP}}}$$

$CP\ 95\%\ CI$ = proportion of 95% confidence intervals around the effect estimate that contain the population value

$Power$ = the proportion of times the 95% confidence intervals does not contain zero, when the true population value is nonzero

3. Methodological Research

Simulation Study - Results

UNIVERSITY AT ALBANY
State University of New York

<u>Relative bias</u>

- Model 0:
    - The number of cases has a large but non-statistically significant effect on the relative bias for estimating the intervention effect, $F(2, 7) = 1.30$, $p = .3304$, $\eta_p^2 = 0.2536$.
    - Intervention effect estimate is unbiased (relative bias < .05) in all conditions.

- Models 1-3:
    - None of the design factors have a statistically significant (i.e., $p < .001$) and large effect (i.e., $\eta_p^2 \geq 0.14$) on the relative bias.
    - In general, intervention and moderator effects estimates (i.e., gender and study quality) are biased (i.e., relative bias > .05) only in conditions with 10 studies and 4 cases.

3. Methodological Research
Simulation Study - Results

UNIVERSITY AT ALBANY
State University of New York

Mean Squared Error (MSE)

- Models 0-3:
  - The number of studies has a statistically significant (i.e., p < .0.01) and large effects (i.e., $\eta_p^2 \geq 0.14$) on the MSE of the intervention/moderator effect estimates. The larger the number of studies (independent of other parameter values), the smaller the MSE.
- Models 1-3:
  - The number of cases has a statistically significant (i.e., p < .001) and large effect (i.e., $\eta_p^2 \geq 0.14$) on the MSE for estimating level-2 moderator effects.

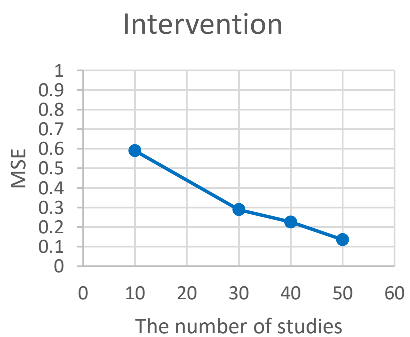
3. Methodological Research

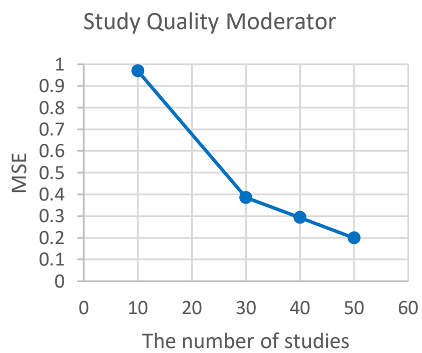
Simulation Study - Results

<u>MSE – Model 1</u>

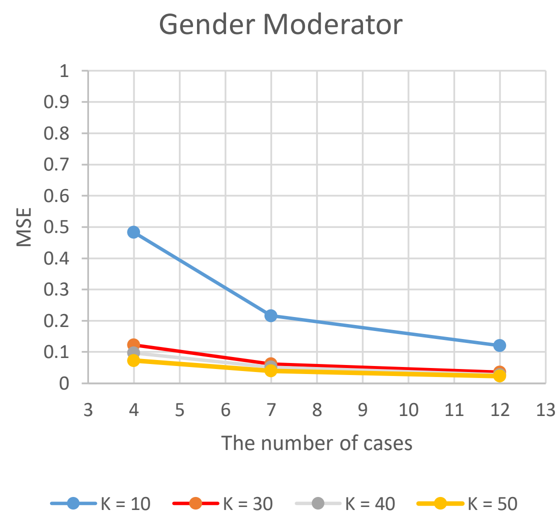
$[\gamma_{100} = 2, \gamma_{110} = 2, \gamma_{101} = 2, I = 20]$

MSE as a function of number of studies [*J = 12*]





MSE as a function of number of studies/cases



3. Methodological Research

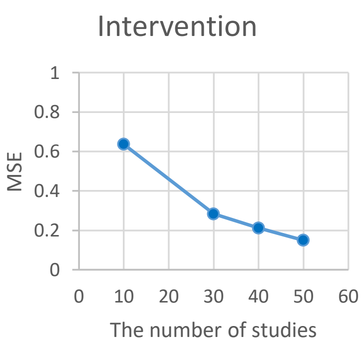
Simulation Study - Results

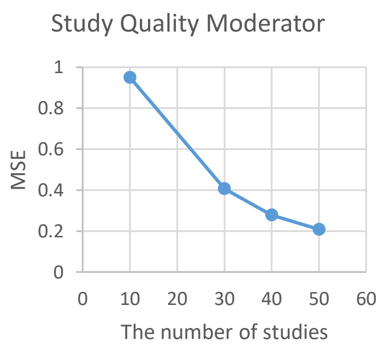
UNIVERSITY AT ALBANY
State University of New York

MSE – Model 2

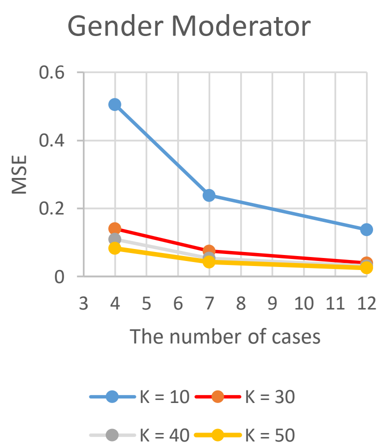
$[\gamma_{100} = 2, \gamma_{110} = 2, \gamma_{120} = 0.5, \gamma_{101} = 2, I = 20]$

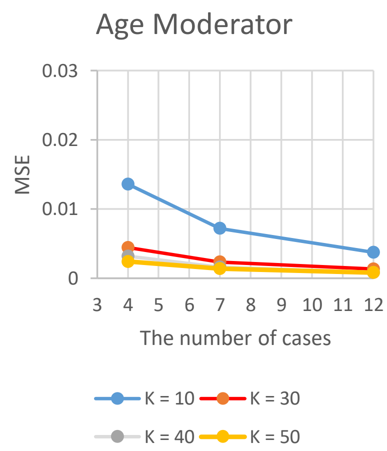
MSE as a function of number of studies
[J = 12]





MSE as a function of number of studies/cases



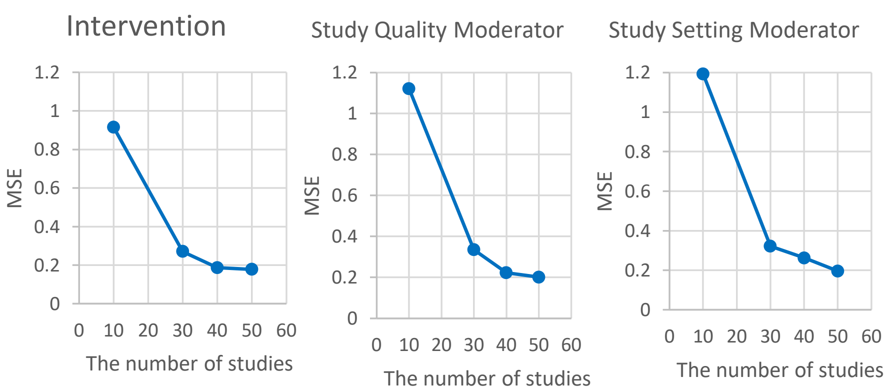


3. Methodological Research
Simulation Study - Results

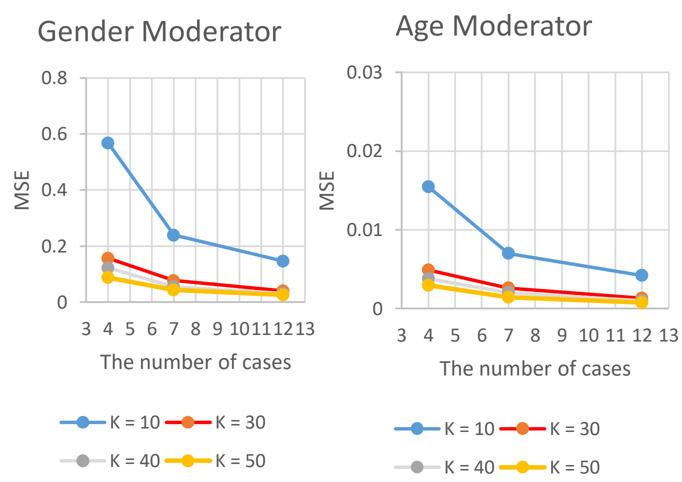
UNIVERSITY AT ALBANY
State University of New York

MSE – Model 3
$[\gamma_{100} = 2, \gamma_{110} = 2, \gamma_{120} = 0.5, \gamma_{102} = 2, \gamma_{101} = 2, I = 20]$

MSE as a function of number of studies
[$J = 12$]



MSE as a function of number of studies/cases



3. Methodological Research

Simulation Study - Results

Relative Standard Error Bias

- Models 0-3:
  - The number of studies has a statistically significant (i.e., $p < .001$) and large effect (i.e., $\eta_p^2 \geq 0.14$) on relative standard error of the intervention/moderator effect estimates (except for estimating age moderator effect)
  - In general, the absolute values of relative standard error are larger than .10 only when the number of studies is 10 and the number of cases is 4.

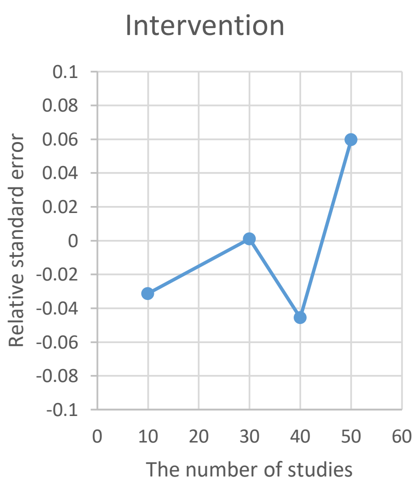
3. Methodological Research

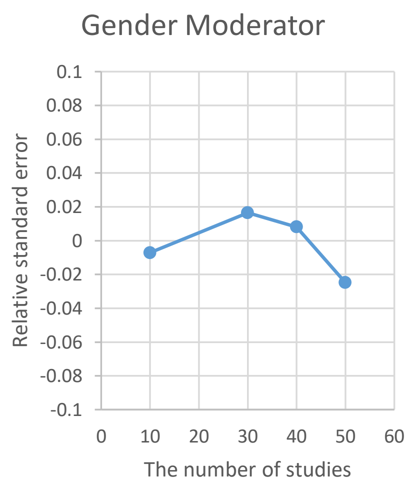
Simulation Study - Results

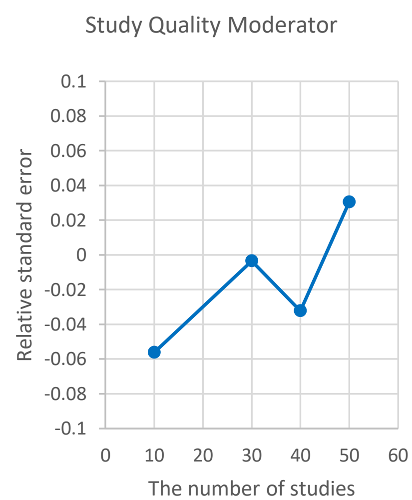
Relative Standard Error Bias – Model 1

$[\gamma_{100} = 2, \gamma_{110} = 2, \gamma_{101} = 2, J = 12\ , I = 20]$

Relative standard error bias as a function of number of studies







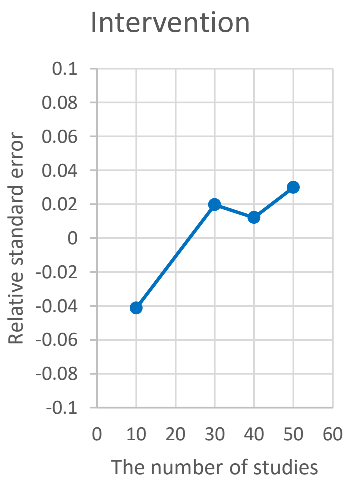
3. Methodological Research

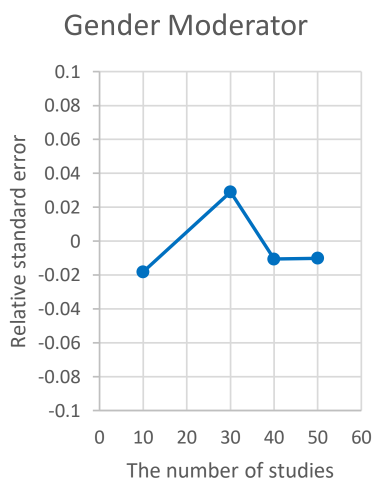
Simulation Study - Results

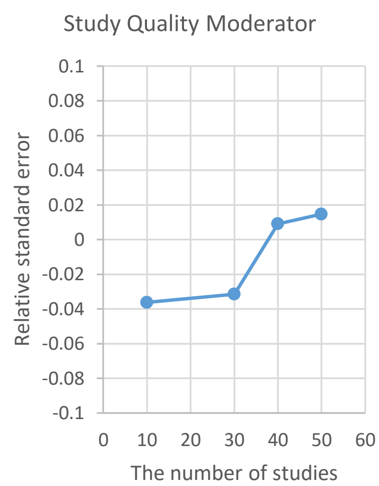
<u>Relative Standard Error Bias – Model 2</u>

$[\gamma_{100} = 2, \gamma_{110} = 2, \gamma_{120} = 0.5, \gamma_{101} = 2,\ J = 12, I = 20]$

Relative standard error bias as a function of number of studies







3. Methodological Research

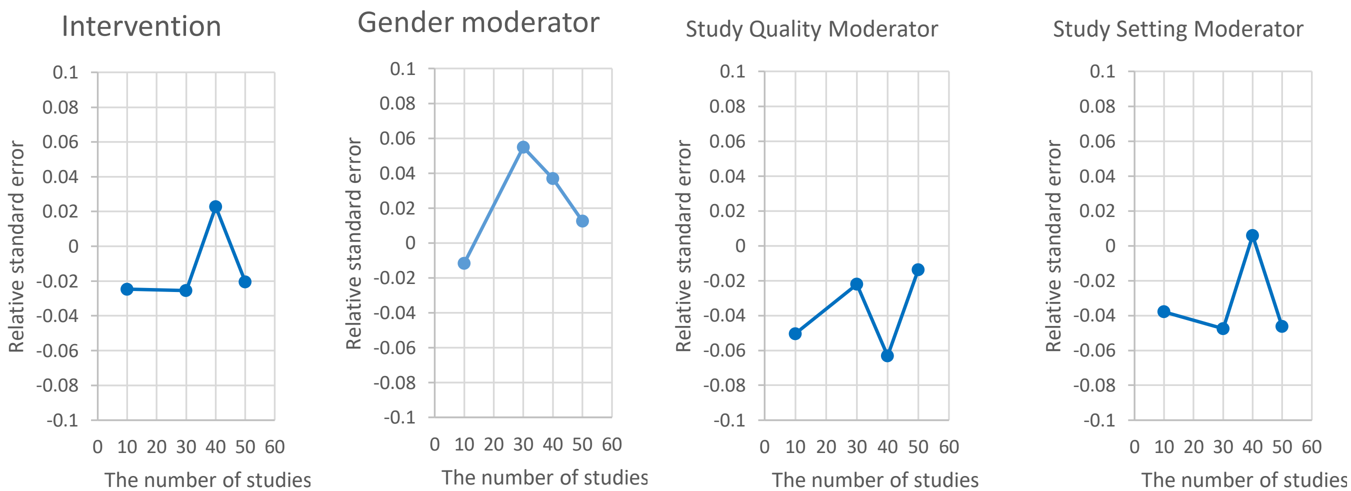
Simulation Study - Results

UNIVERSITY AT ALBANY
State University of New York

Relative Standard Error Bias – Model 3

$[\gamma_{100} = 2, \gamma_{110} = 2, \gamma_{120} = 0.5, \gamma_{101} = 2, \gamma_{102} = 2, J = 12, I = 20]$

Relative standard error bias as a function of number of studies



3. Methodological Research

Simulation Study - Results

UNIVERSITY AT ALBANY
State University of New York

Coverage Proportion of the 95% CI (CP95%)

- Models 0-3:
  - The number of studies only has a statistically significant (p < .001) and large effect (i.e., $\eta_p^2 \geq 0.14$) on the CP95% of the study setting moderator effect estimates in Model 3.
  - For the majority of cases, the CP95% falls within or is very closed to the acceptable range (i.e., 0.93~0.97).
  - Only in Model 3, CP95% of the study setting moderator effect estimates is much smaller than 0.93 (the smallest one is .1440, and CP95% is less than .90 in 3310 conditions). This will be further explored.

3. Methodological Research

Simulation Study - Results

UNIVERSITY AT ALBANY
State University of New York

Power

- Models 0-3:
  - The number of studies has a statistically significant ($p < .001$) and large effect (i.e., $\eta_p^2 \geq 0.14$) on power of the intervention/moderator effect estimates.
  - Magnitude of gender has a statistically significant and large effect on gender moderator effect estimates; magnitude of study quality has a statistically significant and large effect on study quality moderator effect estimates; magnitude of study setting has a statistically significant and large effect on study setting moderator effect estimates.
  - In general, the values of power are less than .80 when number of studies = 10, number of cases = 4, and number of observations = 20

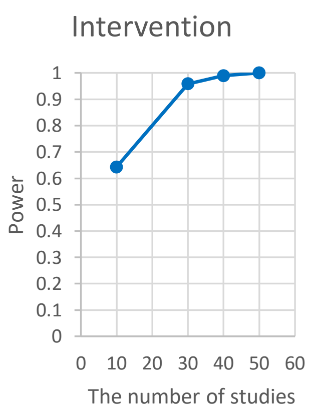
3. Methodological Research

Simulation Study - Results

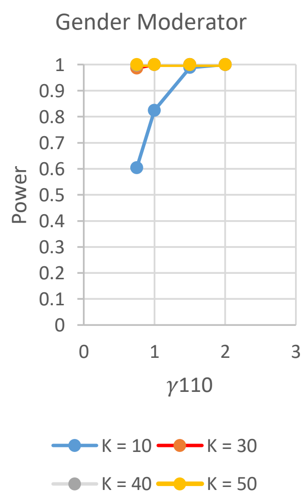
Power – Model 1

$[\gamma_{100} = 2, J = 12, I = 20]$

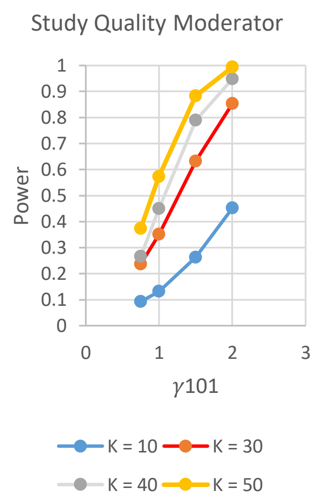
Power as a function of number of studies $[\gamma_{110} = 2, \gamma_{101} = 2]$



Power as a function of number of studies and Gender moderator effect $[\gamma_{101} = 2]$



Power as a function of number of studies and Study quality moderator effect $[\gamma_{110} = 2]$



3. Methodological Research

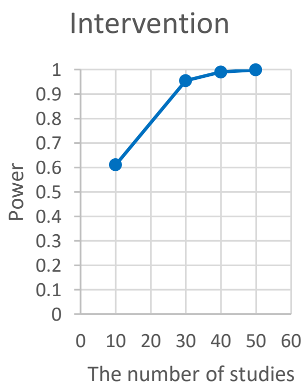
Simulation Study - Results

UNIVERSITY AT ALBANY
State University of New York

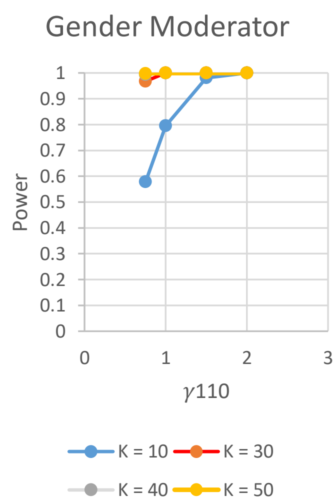
Power – Model 2

$[\gamma_{100} = 2, \gamma_{120} = 0.5, J = 12, I = 20]$

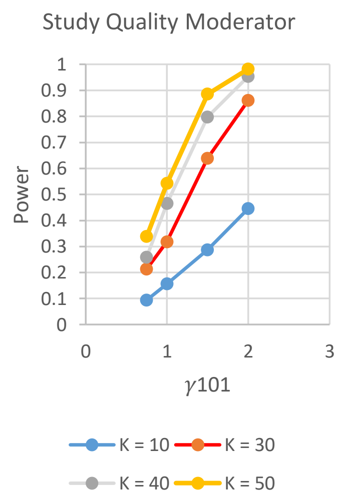
Power as a function of number of studies $[\gamma_{110} = 2, \gamma_{101} = 2]$



Power as a function of number of studies and Gender moderator effect $[\gamma_{101} = 2]$



Power as a function of number of studies and Study quality moderator effect $[\gamma_{110} = 2]$



3. Methodological Research

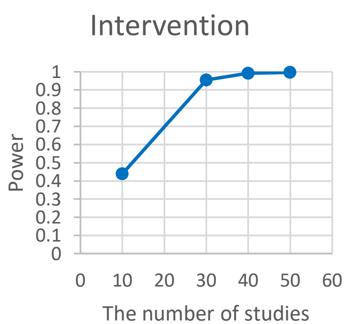
Simulation Study - Results

UNIVERSITY AT ALBANY
State University of New York

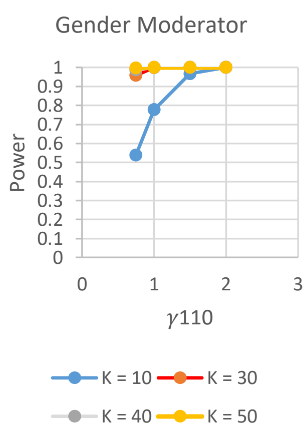
Power – Model 3

$[\gamma_{100} = 2, \gamma_{120} = 0.5, J = 12, I = 20]$

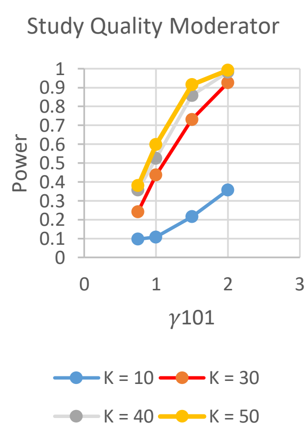
Power as a function of number of studies
$[\gamma_{110} = 2, \gamma_{101} = 2, \gamma_{102} = 2]$



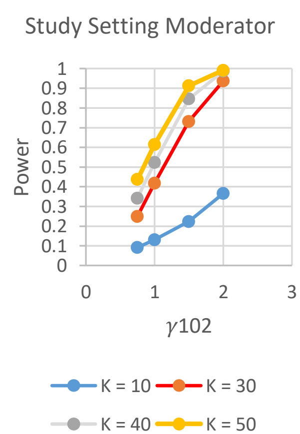
Power as a function of number of studies and Gender moderator effect
$[\gamma_{101} = 2, \gamma_{102} = 2]$



Power as a function of number of studies and Study quality moderator effect
$[\gamma_{110} = 2 , \gamma_{102} = 2]$



Power as a function of number of studies and Study setting moderator effect
$[\gamma_{110} = 2 , \gamma_{101} = 2]$



4. Conclusion

UNIVERSITY AT ALBANY
State University of New York

- The number of studies has an impact on the statistical properties of the intervention and moderator effect estimates across the models.
- In addition, the number of cases has an impact on the statistical properties of the moderator effects on levels 2 and levels 3.
- Consistent with past research, unit changes made at the third level (i.e., number of studies) and the second level (i.e., number of cases) of the hierarchical linear model have a larger effect on estimates compared to units at the lower level (number of measurement occasions).
- We do not recommend using IPD meta-analysis, with the inclusion of moderators, when the number of studies is small ($k = 10$). When number of studies is large ($k \geq 30$), the statistical properties of intervention and moderator effect estimates are appropriate, regardless of the number of cases, number of measurement occasions, and the magnitude of intervention and moderator effects.

4. Future Research

UNIVERSITY AT ALBANY
State University of New York

<u>Many recent developments in the field. Here are just a few examples:</u>

- Further development of the PowerSCED shiny tool

  Xu, X., Moeyaert, M., & Yang, Y. (2021). PowerSCED (Version 1.0) [Web application]. https://xinyunxu.shinyapps.io/PowerSCED/_w_8b4d5ac0

- Bayesian estimation procedure

  Follow-up to: Moeyaert, M., Rindskopf, D., Onghena, P., & Van den Noortgate, W. (2017). Multilevel modeling of single-case data: A comparison of Maximum Likelihood and Bayesian estimation. *Psychological Methods*, *22*(4), 760-778. doi:10.1037/met0000136

- Bayesian mediation analysis

  Follow-up to: Miočević, M., Klaassen, F., Geuke, G., **Moeyaert, M.,** & Maric, M. (2020). Using Bayesian methods to test mediators of intervention outcomes in single case experimental designs.*Evidence-based Communication Assessment and Intervention, 14(1-2), 52-68.* https://doi.org/10.1080/17489539.2020.1732029

- Weighting strategies for multilevel meta-analysis of single-case experimental design data with multiple outcomes.
- New WWC standards (comparison with older versions)

UNIVERSITY AT ALBANY
State University of New York

Usage of IPD Meta-Analysis for Synthesis of SCD Studies

Questions?

Mariola Moeyaert
mmoeyaert@albany.edu
Associate Professor of applied Statistics
University at Albany – SUNY
September 30, 2022