

# Characteristics of Moderators in Meta-Analyses of Single-Case Experimental Design Studies

Behavior Modification

1–36

© The Author(s) 2021

Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/01454455211002111

[journals.sagepub.com/home/bmo](https://journals.sagepub.com/home/bmo)

Mariola Moeyaert<sup>1</sup> , Panpan Yang<sup>1</sup>,  
Xinyun Xu<sup>1</sup>, and Esther Kim<sup>1</sup>

## Abstract

Hierarchical linear modeling (HLM) has been recommended as a meta-analytic technique for the quantitative synthesis of single-case experimental design (SCED) studies. The HLM approach is flexible and can model a variety of different SCED data complexities, such as intervention heterogeneity. A major advantage of using HLM is that participant and-or study characteristics can be incorporated in the model in an attempt to explain intervention heterogeneity. The inclusion of moderators in the context of meta-analysis of SCED studies did not yet receive attention and is in need of methodological research. Prior to extending methodological work validating the hierarchical linear model including moderators at the different levels, an overview of characteristics of moderators typically encountered in the field is needed. This will inform design conditions to be embedded in future methodological studies and ensure that these conditions are realistic and representative for the field of SCED meta-analyses. This study presents the results of systematic review of SCED meta-analyses, with the particular focus on moderator characteristic. The initial search yielded a total of 910

<sup>1</sup>University at Albany—SUNY, NY, USA

## Corresponding Author:

Mariola Moeyaert, School of Education, Department of Educational and Counseling Psychology, Division of Educational Psychology & Methodology, The University at Albany—SUNY, 1400 Washington Avenue, Albany, NY 12222, USA.

Email: [mmoeyaert@albany.edu](mailto:mmoeyaert@albany.edu)

articles and book chapters. After excluding duplicate studies and non peer-reviewed studies, 658 unique peer-reviewed studies were maintained and screened by two independent researchers. Sixty articles met the inclusion criteria and were eligible for data retrieval. The results of the analysis of moderator characteristics retrieved from these 60 meta-analyses are presented. The first part of the results section contains an overview of moderator characteristics per moderator level (within-participant level, participant level, and study level), including the types of moderators, the ratio of the number of moderators relative to the number of units at that level, the measurement scale, and the degree of missing data. The second part of the results section focuses on the metric used to quantify moderator effectiveness and the analysis approach. Based on the results of the systematic review, recommendations are given for conditions to be included in future methodological work.

### **Keywords**

meta-analysis, single-case experimental designs, moderators

Single-case experimental designs (SCEDs) are becoming increasingly popular as a means to establish an evidence base for interventions (Kratochwill & Levin, 2010), especially in behavioral sciences. Along with the increased number of published SCED studies, there is a growing interest in the quantitative synthesis of SCED data across studies (Jamshidi et al., 2018). Meta-analysis can be used to quantitatively summarize SCED studies in a standardized, objective, reliable, and valid manner (Glass, 1976; Kratochwill et al., 2010; Shadish et al., 2013). By synthesizing the effectiveness of an intervention across a large body of literature, an evidence-base on a particular intervention can be created and important decision can be made based on scientific evidence. In contrast to the well-established and broadly applied methods for meta-analysis of group-comparison design studies (e.g., Cohen's *d*, Borenstein et al., 2009; Hedges' *g*, Hedges & Olkin, 1985), there is a lack of consensus about which methods can be applied to meta-analyze single-case experimental design (SCED) studies (Van den Noortgate & Onghena, 2008). Similar to group-comparison design studies, the effect of an intervention can be quantified by comparing the mean of data obtained during an experimental condition and the mean of data obtained during a baseline condition (i.e., no intervention is given; Hedges et al., 2012, 2013). The fundamental difference between a group-comparison design study and an SCED study is the unit of analysis. For group-comparison designs the average score

across participants assigned to the baseline condition is compared to the average score across participants assigned to the experimental condition. In contrast, individual participants are the unit of analysis in SCEDs. Participants in SCED are not assigned to a treatment and control group, but are repeatedly measured during both baseline and intervention conditions. As a consequence, the participant serves as its own control and no comparison group is needed (Kratochwill et al., 2010; What Works Clearinghouse, 2020). In order to make inferences related to intervention effectiveness beyond an individual participant, SCED relies upon within study and across study replication (Horner et al., 2005; Kazdin, 2011). The experiment needs to be replicated across individuals (which is usually accomplished within one SCED study). Moreover, either direct or systematic replication has to occur across multiple studies in order to establish an evidence-base and enhance external validity (Ferron & Scott, 2005; Horner et al., 2005). As a consequence, SCED research is time consuming and demanding. Because of the aforementioned fundamental differences between group-comparison designs and SCEDs, analysis techniques appropriate for group-comparison designs are not transferable to SCED and different methods need to be considered.

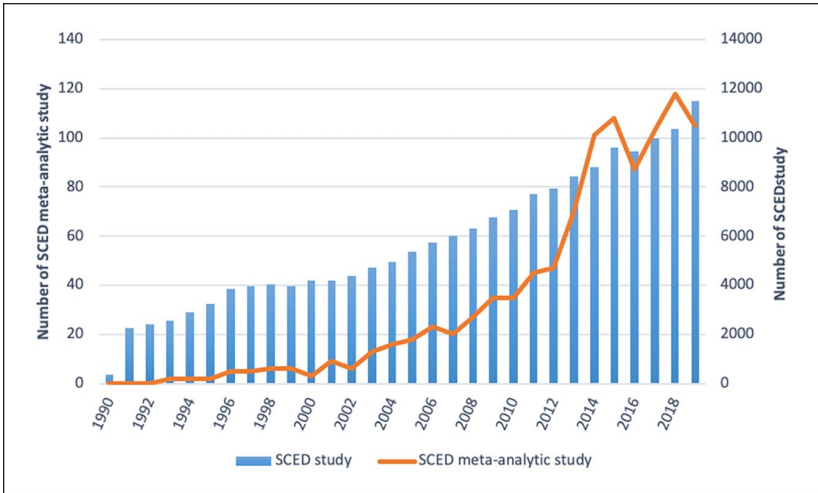
### *Types of Analysis for SCED*

During the last decade, a variety of different methods have been developed to quantify intervention effectiveness within and across SCED studies. Examples include metrics for quantifying the intervention effect per participant (e.g., non-overlap metrics and regression-based metrics, Parker, Vannest, & Davis, 2011) to approaches suitable to summarize intervention effectiveness across participants (Ferron et al., 2009) and even across studies, such as hierarchical linear modeling (Van den Noortgate & Onghena, 2003a, 2003b). The metrics to quantify the intervention effectiveness can be classified into four broad categories: (a) non-overlap metrics (e.g., PND, Scruggs et al., 1987; IRD, Parker et al., 2009; NAP, Parker & Vannest, 2009; Tau, Parker, Vannest, Davis, & Sauber, 2011; Tarlow, 2017), (b) regression based metrics (including hierarchical linear modeling as an extension) (e.g., Moeyaert, Ugille, et al., 2014; Van den Noortgate & Onghena, 2003a, 2003b), (c) log ratio metrics (Pustejovsky, 2015), and (d) standardized mean difference metrics (BC-SMD, Hedges et al., 2012, 2013). Because of the limited generalizability of any one SCED study, across-participant approaches are appealing. Across participant approaches can be used to summarize intervention effectiveness per study allowing to make more generalizable conclusions.

## *Meta-Analysis of SCED using HLM*

One meta-analytic technique that is appropriate for meta-analysis of SCED is hierarchical linear modeling (HLM) as this technique takes the nested data structure into account: repeated measures are nested within participants and participants are nested within studies (Moeyaert, Ferron, et al., 2014; Shadish et al., 2013). The statistical properties of this approach has been extensively studied and empirically validated using large-scale Monte Carlo simulation studies (Ferron et al., 2009; Moeyaert et al., 2013a, 2013b; Ugille et al., 2012). Given its desirable statistical properties, the HLM approach has been recommended for the quantitative synthesis of SCED study results across studies (Shadish et al., 2013). The HLM approach is flexible and can model a variety of different data complexities, such as autocorrelation (Maggin et al., 2011), linear and non-linear trends (Shadish et al., 2013; Van den Noortgate & Onghena, 2003b), count outcomes (Declercq et al., 2019; Shadish et al., 2013), and intervention heterogeneity (Baek & Ferron, 2013). However, one topic that did not yet receive attention and is in need of methodological research is the inclusion of moderators in the context of HLM of SCED studies. The lack of research on moderator analysis is worrisome as applied SCED meta-analyses including moderators have been published without knowing whether the chosen metric and analytic approach is suitable and powerful enough (Heyvaert et al., 2012, 2014; Hurwitz et al., 2015; Stone, 2011; Vanderkerken et al., 2013; Wang et al., 2013).

Previous methodological work in contexts of quantitative synthesis of SCEDs evaluated the statistical properties of the intervention effect estimate across SCED studies (Van den Noortgate & Onghena, 2003a; Zimmerman et al., 2018) and the between-study and between-case variability in these intervention effect estimates (Moeyaert, Ugille, et al., 2014). The multi-level meta-analytic model is one of the statistical analysis techniques that has been used and empirically validated for this purpose (Moeyaert et al., 2013a, 2013b, 2016; Petit-Bois et al., 2016; Ugille et al., 2012). However, no research up to date has focused on the statistical properties of multi-level models that include moderators. By adding moderators at the case and/or study level, the unexplained variability in intervention effects between cases and/or studies can decrease. Prior to extending methodological work validating the multi-level modeling approach including moderators at the different levels, the characteristics of moderators typically encountered in the field is needed. The overview will inform design conditions to be embedded in future methodological work and ensure that these conditions are realistic and representative for the field of SCED meta-analyses. This study presents the results of a systematic review of SCED meta-analyses, with a particular



**Figure 1.** Graphical overview of the number of SCED studies and SCED meta-analyses published between 1990 and 2019.

Note. The online database Web of Sciences was used to create the graph. The SCED studies were searched with following string for topic: TS= (“single case” OR “single-case” OR “single case experiment” OR “single-case experimental design” OR “single case experimental” OR “MBD” OR “multiple-baseline” OR “multiple baselines” OR “N= 1” OR “n= 1” OR “intra subject” OR “intra-subject” OR “interrupted time series” OR “alternating treatment design” OR “reversal design”). The SCED meta-analytic studies were searched with the above string combined with TS= (“meta analysis” OR “meta-analytic” OR “systematic review” OR “meta-analysis” OR “meta-analyses” OR “systematic reviews”). The search is limited to the English language.

focus on moderator characteristics. SCED meta-analyses eligible for inclusion are further categorized into three broad categories: (a) meta-analyses including moderators in the analysis, (b) meta-analyses recognizing and presenting moderators (but did not include the moderators as part of the analysis plan), and (c) meta-analyses lacking moderators. To present a focused and in-depth discussion of SCED moderator characteristics, the focus of this study is on the first category of SCED meta-analyses (i.e., meta-analysis including moderators in the analysis).

### Reviews of SCED Meta-Analyses

Figure 1 gives a graphical overview of the number of SCED studies and SCED meta-analyses published between 1990 and 2019 using the Web of Sciences database. This illustrates that there has been an exponential increase

in the number of SCED studies published over the last three decades. Because more research evidence from primary level SCED studies is available, there is also an exponential increase in the number of meta-analyses (and systematic reviews in general) of SCEDs (see Figure 1 and Jamshidi et al., 2018).

Several studies have been published summarizing methodological aspects and data characteristics of these meta-analyses and systematic reviews (e.g., Beretvas & Chung, 2008; Farmer et al., 2010; Jamshidi et al., 2020; Maggin et al., 2011; Schlosser et al., 2008). Jamshidi et al. (2020) conducted a systematic review of SCED meta-analyses. Their systematic review covered a large timespan (1985–2015), and included general data characteristics and study design characteristics of SCED meta-analyses. In addition, an overview was provided of the kind of analyses done per primary level SCED study and meta-analysis. Jamshidi et al. (2020) found that 130 out of the 173 meta-analyses conducted a moderator analysis. These moderators were intervention and participant characteristics. However, specific details such as the measurement scale and the number of case-specific and intervention-specific moderators per meta-analysis were not reported. The focus of another systematic review conducted by Jamshidi et al. (2018) was on the methodological quality of SCED meta-analyses. Jamshidi and colleagues assessed the methodological quality of 178 SCED meta-analyses published between 1985 and 2015. They used the *Revised-Assessment of Multiple Systematic Reviews* (R-AMSTAR) checklist as a guideline. They found that SCED meta-analytic studies did better on some criteria (e.g., “doing a comprehensive literature search,” “providing the characteristics of the included studies”), while it was not the case for other criteria, such as “reporting an assessment of the likelihood of publication bias” and “using the methods appropriately to combine the findings of studies.” Jamshidi and colleagues concluded that the methodological quality of SCED meta-analyses, in general, was low, although it has slightly increased over time. The characteristics of moderators included in SCEDs meta-analyses were not considered as this was not an item included in the R-AMSTAR checklist.

Other systematic reviews of meta-analyses mainly focused on characteristics of effect size metrics used in meta-analyses. For instance, Maggin et al. (2011) reviewed 68 SCED meta-analyses between 1985 and 2009 focusing on participants with and at-risk for disabilities and found that the percentage of non-overlapping data (PND) was the most frequently used metric, followed by the standardized mean difference (SMD). The mean, weighted mean or median was most frequently used to synthesize the results across studies. Farmer et al. (2010) showed similar results to Maggin et al. (2011). Schlosser et al. (2008) focused on exploring the characteristics of PND by reviewing 45 meta-analytic studies from 1985 to 2008 and found that most

included studies aggregated the scores of PND across different studies. Compared to the median PND, the average PND was more frequently used to represent the overall intervention effect across studies. Besides examining the effect size metric, Beretvas and Chung (2008) reviewed 25 meta-analyses of SCEDs and explored how the dependency in the original studies with multiple interventions, outcomes, and participants was handled in the meta-analyses. They found that most studies did not report how they dealt with the dependency.

Based on the review of previous systematic reviews of SCED meta-analyses presented in previous sections, it can be concluded that summarizing moderator characteristics has not been thoroughly considered. In addition, items related to moderators have not been considered for inclusion in reporting guidelines to assess the quality of SCED meta-analyses and systematic reviews. This stands in contrast to reporting guidelines developed for group-comparison design systematic reviews and meta-analyses (see Table 1). The items related to moderator characteristics included in these guidelines can be considered as a source to develop similar items to be included in meta-analytic quality assessment tools and reporting guidelines such as the R-AMSTAR checklist (see Table 1).

### *SCED Reporting Guidelines and Quality Assessment Tools*

Meta-analyses are depending on the information that is reported in the primary level studies eligible for inclusion in the meta-analysis. Therefore, items related to moderator variables are ideally included in reporting guidelines and quality assessment tools for primary level SCED studies. Lobo et al. (2017) provided an overview of quality assessment and reporting tools available for SCEDs; Quality indicators (Horner et al., 2005); Evaluative method (Reichow et al., 2008); Evidence in Augmentative and Alternative Communication Scales (EVIDAAC, Schlosser et al., 2009); Single-Case Experimental Design (SCED, Tate et al., 2008), Logan et al. scales (Logan et al., 2008), and Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE, Tate et al., 2016). Tate et al. (2008) developed the Single-Case Experimental Design Scale to evaluate the quality of SCED studies. Later on, Tate et al. (2016) further developed the Single-Case Reporting Guideline in BEhavioural Interventions (SCRIBE) to provide a checklist that helps single-case studies to be published as well as journal reviewers and editors to evaluate the quality of single-case studies.

Another source focusing on quality assessment of primary level SCEDs is the What Works Clearinghouse (WWC) technical documentation developed by the Institute of Education Sciences (Kratochwill et al., 2010; What Works

**Table 1.** Reporting Guidelines and Assessment Tools Related to Moderators in the Context of Single-Case Experimental Design (SCED) Studies and Group-Comparison Design Systematic Reviews and Meta-Analyses.

Design	Reporting guidelines/ assessment tools	Included items/ categories	Items/categories related to moderators
Group-comparison design systematic reviews and meta-analyses	Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Moher et al., 2009)	27 Items	Two items are related to moderators For the results of study characteristics, it requires a meta-analytic study reporting the data characteristics for each included study (e.g., characteristics of study size, participants, interventions, comparisons, outcomes, study design, and follow-up period) For the results of individual studies, it requires to present (a) simple summary data for each intervention group and (b) effect estimates and confidence intervals for each primary study
	Assessment of Multiple Systematic Reviews (AMSTAR) (Shea et al., 2007)	11 Items	Two items are related to moderators One item evaluates whether the characteristics of the included studies were provided on the participants, interventions, and outcomes in an aggregated form. All the analyzed characteristics (e.g., age, race, gender, relevant socioeconomic data, disease status, duration, severity, etc.) in the original studies should be reported. The other item evaluates whether the scientific quality of the included studies was assessed and documented, which refers to the moderator of validity of included studies.
	Revised Assessment of Multiple Systematic Reviews (R-AMSTAR) (Kung et al., 2010)	11 Items	Same two items as AMSTAR are related to moderators

(continued)



**Table 1. (continued)**

Design	Reporting guidelines/ assessment tools	Included items/ categories	Items/categories related to moderators
SCED studies	Single-Case Experimental Design Scale (SCED Scale) (Tate et al., 2008)	11 Items	Five items are related to moderators One item evaluates clinical history of participants, which includes participant's demographic and injury characteristics The second related item evaluates whether a SCED study identifies a precise, repeatable and operationally defined target behavior The third related item refers to study design of a SCED study The fourth related item refers to inter-rater reliability The last related item evaluates the generalization of a SCED study
	Single-Case Reporting Guideline in Behavioural Interventions (SCRIBE) (Tate et al., 2016)	26 Items	Five items are related to moderators One item refers to the guideline for reporting study design The second related one is about how to report participant(s) or unit(s) (i.e., reporting the selection criteria and participant characteristics) The third related one is related to how to report context, which is the characteristics of the setting and location where the study was conducted The fourth related one refers to target behaviors and outcome measures (e.g., whether the target behaviors and outcome measures are well operationally defined) The fifth related one involves the procedure of intervention and procedural fidelity
	Logan et al. scales (Logan et al., 2008)	14 Items	Two items are related to moderators, which are study design and description of participants and settings

(continued)

**Table 1. (continued)**

Design	Reporting guidelines/ assessment tools	Included items/ categories	Items/categories related to moderators
	Evidence in Augmentative and Alternative Communication Scales (EVIDAAC) (Schlosser et al., 2009) Quality indicators (Horner et al., 2005)	19 Items  Seven categories	Four items are related to moderators, which are participant characteristics, physical setting, interobserver agreement, and treatment integrity  Three categories are related to moderators The first category refers to the guideline for description of participants and settings, such as participant's age, gender, disability, the physical setting, etc. The second and third related ones are both related to validity of a SCED study. One is external validity and the other one is social validity.
	Evaluative method (Reichow et al., 2008)	Two broad categories (total 12 detailed indicators)	Five indicator is related to moderators: One is about participant characteristics The second related one refers to interobserver agreement The remaining three related indicators involve fidelity, generalization and/or maintenance, and the social validity of a SCED study
	What Works Clearinghouse Standards (WWW) (Kratochwill et al., 2010) Council for Exceptional Children: Standards for evidence-based practices in special education (Council for Exceptional Children [CEC], 2014) <sup>a</sup>	Five categories  Eight categories	One category is related to moderators, which is <i>interassessor agreement</i>  Six categories are related to moderators, which are (1) context and setting (e.g., type of school, geographic location), (2) participants characteristics (e.g., age, gender, disability status), (3) intervention agent (e.g., role of the intervention agent, whether having any specific training), (4) implementation fidelity, (5) internal validity, (6) outcome measures and target behaviors

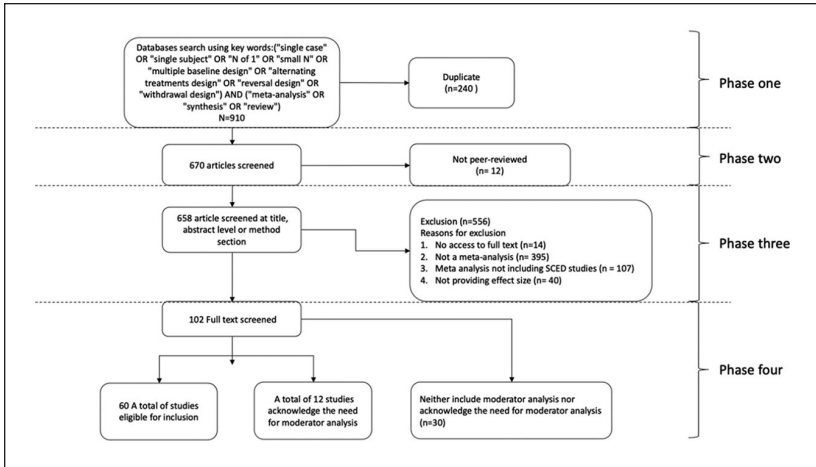
<sup>a</sup>CEC can be used for both group-comparison designs and SCEDs.

Clearinghouse [WWC], 2020). Specifically, when reviewing SCEDs, the study rating criteria of WWC can be used to categorize the SCED studies into three levels of quality, namely *Meets WWC SCD Standards Without Reservations*, *Meets WWC SCD Standards With Reservations*, or *Does Not Meet WWC SCD Standards*. Other available quality assessment tools for SCEDs, such as quality indicators from Horner et al. (2005), evaluative method from Reichow et al. (2008), and Logan et al. scales (Logan et al., 2008), focus on similar aspects as Tate et al. (2008, 2016) to assess the quality of SCEDs (for more detail, review Horner et al., 2005; Reichow et al., 2008).

As we can be deduced from Table 1, existing checklists for quality assessment of group-comparison design meta-analyses and primary level SCEDs, the characteristics of moderators are embedded to some extent (CEC, 2014; Kratochwill et al., 2010; Kung et al., 2010; Moher et al., 2009; Shea et al., 2007; Tate et al., 2008, 2016; WWC, 2020). However, they do not mention what specific characteristics of moderators need to be reported (e.g., measurement scale of moderators and modeling approach). Information related to moderator characteristics of SCED studies are important not only to determine which analytic methods are best used in meta-analyses (Jamshidi et al., 2020), but also to inform the design conditions and parameter values to be included in future Monte-Carlo simulation studies. Including moderators can help explain variability in intervention effectiveness between individuals and-or studies.

### *Purpose and Research Questions*

The purpose of this systematic review of SCED meta-analyses is to provide insights into moderator characteristics. First, the review is designed to provide an overview of the type of moderators studied in SCED meta-analyses. To enhance the discussion, moderators are classified within three “levels” (a) the outcome and intervention level (i.e., level 1, related to within-participant characteristics), (b) the participant level (i.e., level 2), and (c) the study level (i.e., level 3). Per moderator level, an overview and description of commonly encountered moderators is provided. Numerics related to the number of moderators at these levels and the ratio of the number of moderators at a certain level relative to the number of units at that level are reported. In addition, the measurement scale of the moderators at the different levels is discussed (i.e., Nominal, Ordinal or Continuous [Interval/Ratio]). Information related to the degree of missing data per moderator is captured as well. In sum, the first section of the results contains key information related to the moderator characteristics, presented per level. Second, in addition to these specific moderator characteristics, aspects specific to the moderator analysis are captured and



**Figure 2.** Flowchart of the systematic review of single-case experimental design meta-analyses.

discussed, namely (a) the metric used to quantify moderator effectiveness (i.e., non-overlap metrics, regression based metrics, log ratio metrics, and standardized mean difference metrics) (b) the unit of analysis (participant-specific, study-specific or across studies) and (c) the specific approach used to combine metrics across cases and/or across studies (i.e., description in words, frequency table, quantitative metric or statistical modeling).

## Methods

### Systematic Literature Search

The following six online databases were used to conduct the systematic search: PsycINFO, Web of Science, Science Direct, Medline PubMed, ERIC, and CINAH. The systematic literature search procedure as outlined by Jamshidi et al. (2020) was replicated in current research (by the two independent coders) as this is the most extended and profound systematic review of SCED meta-analyses identified. To be included in the systematic review, the study is required to be (a) available in English, (b) peer-reviewed with full-text availability, (c) published between 2016 and 2019, and (d) a meta-analysis including SCED studies. In addition, the study must (e) provide an effect size, and (f) include moderator analysis. These inclusion and exclusion criteria are displayed in Figure 2. As suggested by Jamshidi, two sets of search

strings are specified in all databases: (“single case” OR “single subject” OR “N of 1” OR “small N” OR “multiple baseline design” OR “alternating treatments design” OR “reversal design” OR “withdrawal design”) AND (“meta-analysis” OR “synthesis” OR “review”). In order to ensure that the same search procedure was applied as in Jamshidi et al. (2020), the two independent researchers replicated Jamshidi’s search procedure for 1985 to 2015 for all of the six databases and verified whether the same number of studies per database were retained. Discrepancies between Jamshidi and the two independent researchers were found for all databases. Therefore, Jamshidi was contacted so that these discrepancies could be resolved. Once the discrepancies were resolved, the search procedure together with the options per database were refined. Discrepancies between Jamshidi and the two independent researchers were found for certain databases, namely Web of Science, ERIC and CINAHL. Different access to use different online reference systems such as EBSCOhost was the cause of the discrepancies. The specific search procedure per database is presented in Supplemental Appendix B as search options slightly varied per database. By outlining these details, the search can be replicated and the same number of SCED meta-analyses eligible for inclusion should be identified. The average IOA for the number of retrieved studies across the six databases is 84%. The database specific IOA in terms of the number of studies that were initially identified for each database (i.e., PsycINFO, Web of Science, Science Direct, Medline PubMed, ERIC, and CINAHL) are 93, 99,49, 100; 71, and 89%, respectively.

The same two doctoral students applied the inclusion and exclusion criteria independently and resolved any discrepancies in the number of excluded studies per criterion. The proportion of agreement at phase one (i.e., exclusion of duplicates, see Figure 2) was 99%. The reason for the discrepancy was that a proportion of the included studies were duplicates with slightly different titles. After reviewing the articles with similar titles, discrepancies were resolved. The proportion of agreement at phase two (i.e., exclusion of non-peer reviewed articles, see Figure 2) was 17%. The reason for this large degree of discrepancy was that the two doctoral students used slightly different strategies in identifying peer-review journals. After applying the same strategy, agreement was reached. The proportion of agreement at phase three (i.e., exclusion of articles that were not SCED meta-analyses and do not present an effect size metric, see Figure 2) was 92%. Phase three had four exclusion reasons (see Figure 2), and one doctoral student identified some articles fitting more than one of these exclusion reasons. For example, one article was identified to be a meta-analysis of group designs. In addition, this meta-analysis focused on the quality and characteristics of the included studies without providing an effect size. This double counting of exclusion reason caused the

discrepancy. After discussing how to label the reason for exclusion for each moderator, 100% agreement was obtained. The proportion of agreement at phase four (i.e., exclusion of articles that did not include moderators or recommended or acknowledge moderator analysis, see Figure 2) was 78%. The disagreement was because one of the doctoral students included the articles if they emphasized the importance of moderators or recommended the moderator analysis for further research. The other doctoral student included articles that acknowledged the possibility of moderator analysis. The discrepancy was solved after the two doctoral students clarified that only articles SCED meta-analyses conducting moderator analysis, recommending moderator analysis or acknowledged the need of moderator analysis (with reporting of the characteristics of the potential moderators) were eligible for inclusion.

### *Independent Coders and Interobserver Agreement*

The systematic literature search for SCED meta-analyses was conducted by two independent researchers enrolled in the doctoral program *Educational Psychology and Methodology*. Both researchers successfully completed an SCED class and/or conducted research apprenticeships related to the methodology of SCEDs. Both researchers have profound expertise in the design and analysis of SCED studies. They can identify SCED studies, differentiate and code different types of SCEDs, and are able to retrieve raw data from primary and meta-analytic SCED studies. The independent researchers conducted a systematic literature search using six online databases. The interobserver agreement (IOA) between the two independent researchers was calculated per database by dividing the number of studies identified by both researchers (i.e., the number of agreements) by the sum of the number of agreements and disagreements. Next, the average IOA across the six databases was calculated.

The data retrieval and coding of SCED meta-analysis was performed by three independent researchers. Two of these researchers were part of the systematic literature search. The third independent researcher was also a doctoral student within the program *Educational Psychology and Methodology* and has a similar background as the other two independent researchers. Given that there are no reporting guidelines for SCED meta-analyses, there is a discrepancy in the way moderator characteristics are reported in SCED meta-analyses. Because of this complexity, the three researchers were first trained using the codebook (i.e., training manual, discussed later). During this training procedure, the three coders independently retrieved data from one SCED meta-analysis. The coding results were compared to identify discrepancies. All discrepancies were discussed until all

coders obtained complete agreement. The coders repeated this procedure for other SCED meta-analyses until no discrepancies were identified. Three rounds of coding training was needed to accomplish this. Upon the completion of the training, the researchers coded independently a fourth SCED meta-analysis and the IOA between the three researchers was calculated. The IOA between the three independent researchers was calculated by averaging the 3% of agreement between the pairs of independent coders (i.e.,  $IOA_{\text{coder1,2,3}} = [IOA_{\text{coder1,2}} + IOA_{\text{coder1,3}} + IOA_{\text{coder2,3}}]/3$ ). The independent researchers coded several variables, including the specific moderators, the level of the moderators, (i.e., level 1, representing the within-participant level; level 2, representing the participant level; and level 3, representing the study level), the number of moderators and units per level, the ratio of the number of moderators relative to the number of units at that level, the measurement scale per moderator, the degree of missing data per moderator, the metric used to quantify moderator effectiveness, and the specific analytic approach used to combine metrics across cases and/or across studies. The IOA for coding of all these major variables of interest was calculated. A list of these variables of interest together with all other coded variables (of secondary interest) is presented in Supplemental Appendix A.

### *Single-Case Experimental Design Studies Eligible for Inclusion*

After applying the exclusion and inclusion criteria, a total of 60 articles were eligible for inclusion. The complete list of these articles is provided in Supplemental Appendix C. The initial search yielded a total of 910 articles and book chapters. After excluding duplicate studies and non-peer-reviewed studies, 658 unique peer-reviewed studies were maintained. All these articles were screened at the title, abstract, and methods level. A total of 556 articles were excluded; 395 because not being a meta-analysis, 107 because not including SCED studies, 40 because the SCED meta-analyses did not include effect sizes; and 14 because no access to the full text was obtained. The full texts of the remaining 102 articles were further screened, and the two doctoral students identified 60 articles that were eligible for final review and coding.

### *Data Retrieval and Coding*

The codebook was created based on the research aim and research questions to ensure that all relevant variables are captured and coded accordingly. Specifically, two parts of information that reveal the characteristics of moderator and the aspect regarding analysis are coded (see Supplemental

Appendix A). Part one of the codebook includes the name, level and measurement scale of the moderator, and ratio between the number of moderators relative to the number of units at each level. Part two includes the metric used to quantify moderator effectiveness and the specific approach used to combine metrics across cases and/or across studies. The complete codebook specifying all coded variables and categories within variables can be found in Supplemental Appendix A. The codebook was used to train the three independent researchers.

The three independent researchers obtained an IOA of 100% for coding the total number of primary studies and participants per SCED meta-analysis. The IOA for moderator identification equaled 78.8%. After identifying the moderator, the three researchers also independently coded the level of each moderator (i.e., the outcome and intervention level, the participant level, and the study level) for which an IOA of 86.6% across the three researchers was obtained. The three researchers obtained complete agreement for coding of the measurement scale of moderators, the metric used to quantify moderator effectiveness and the specific approach used to combine metrics across cases and/or across studies. The degree of missing data of the moderators was also captured and for this an IOA of 86.6% was obtained. The overall IOA between three independent researchers across all the coded variables was 93%. After the three independent researchers were trained using the codebook and the coding reliability was evaluated, each researcher was assigned a set of 20 SCED meta-analyses to code.

## *Data Analysis*

The statistical software program SAS 9.4 (SAS Institute Inc., 2013) was used for quantitative analysis and synthesis. The raw data retrieved from the SCED meta-analyses was entered in separate Excel Sheets by the three independent coders. The excel sheets from the three independent coders were merged per moderator level and imported into SAS 9.4. As there are three moderator levels, (i.e., study level, participant level and within-participant level [i.e., intervention and outcome moderators]), three separate sheets, merged within one big dataset, were created. First, descriptive analyses were run to summarize moderator characteristics per level (i.e., frequency, measurement scale, and information related to the degree of missing data per moderator). Second, SCED metrics (i.e., non-overlap, regression based/HLM, log ratio, and SMD) and analysis approaches (i.e., description in words, frequency table, and quantitative metric or statistical analysis). The results section presents the characteristics of moderators included in at least five SCED meta-analyses. The results for other moderators included in less than five SCED meta-analyses are provided in Supplemental Appendix D.



## Results

A total of 60 SCED meta-analyses met the inclusion criteria and were eligible for data extraction. The results of the analysis of moderator characteristics retrieved from these 60 meta-analyses are presented. The first part of this results section provides an overview of moderator characteristics organized per moderator level, including the types of moderators, the ratio of the number of moderators relative to the number of units at that level, the measurement scale, and information related to the degree of missing data per moderator. The second part of the results section focuses on metrics used to quantify moderator effectiveness and details related to the analysis approach (e.g., description in words, inclusion of frequency or frequency table, reporting of quantitative metric or including a statistical analysis).

### *Moderator Characteristics*

#### *Study level moderators*

*Ratio and type of moderators.* The average ratio between the number of moderators at the study level and the corresponding number of units at that level is 0.22. This means that, on average, meta-analyses with 10 primary SCED studies typically include two study level moderators. This ratio ranges from 0 to 1.40. A total of 24 unique study level moderators (across all eligible SCED meta-analyses) are identified among which 13 are included by at least five meta-analytic studies. These 13 moderators represent commonly reported study level moderators and characteristics of these are summarized in Table 2. An overview of the other, less commonly reported study level moderators, can be found in Supplemental Appendix D. Forty-two out of the sixty SCED meta-analyses include SCED design type as a moderator. Specifically, a total of eight different study designs are identified and summarized in Table 3. The multiple-baseline design is the most commonly reported study design, followed by the reversal design. Changing criterion design is the least popular one. Thirty-eight out of the sixty SCED meta-analyses examine the physical setting the intervention took place in as a moderator. The physical setting includes classroom, home, clinic, community center, playground, and others. Twenty-six studies include SCED quality design standards as a moderator, which comprises either the WWC standards or the CEC standards (both discussed under the section reporting guidelines and quality assessment tools for SCEDs). Furthermore, 23 of these 26 studies discuss the degree to which the WWC standards or CEC standards are met (i.e., fully meet the standards, partly meet the standards or do not meet the standards). Fifteen SCED meta-analyses discuss the interobserver agreement. The maintenance and generalization

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

Moderator level	Moderator	Number (%) <sup>a</sup>	Measurement scale	Missing data <sup>b</sup>	Analysis approach (%) <sup>c</sup>
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 repertoires	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)

**Table 2. (continued)**

Moderator level	Moderator	Number (%) <sup>a</sup>	Measurement scale	Missing data <sup>b</sup>	Analysis approach (%) <sup>c</sup>
Within-participant level	Intervention program <sup>d</sup>	44 (73)	Nominal	No (n=26); yes (n=7); not clear (n=11)	D (98); F (91); Q (73); S (50)
	Intervention agent <sup>d</sup>	28 (47)	Nominal	No (n=13); yes (n=8); not clear (n=7)	D (96); F (100); Q (57); S (43)
	Intervention techniques <sup>d</sup>	21 (35)	Nominal	No (n=13); yes (n=3); not clear (n=5)	D (100); F (100); Q (62); S (43)
	Intervention dosage <sup>d</sup>	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 <sup>d</sup>	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 <sup>d</sup>	7 (12)	Nominal	No (n=4); yes (n=1); not clear (n=2)	D (86); F (71); Q (43); S (14)
	Outcome domain <sup>e</sup>	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 <sup>e</sup>	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).

<sup>a</sup>Percent of SCED meta-analytic studies including this moderator.

<sup>b</sup>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.

<sup>c</sup>Percent of SCED meta-analytic studies using different types of analysis approach.

<sup>d</sup>Intervention specific moderators.

<sup>e</sup>Outcome specific moderators.

**Table 3.** Types of Research Design Included in Meta-Analytic Studies.

Type of study design	Number (%) of meta-analyses using this study design
Multiple baseline	35 (27)
Reversal design	24 (18)
Alternative treatments	19 (15)
AB/ABA	14 (11)
Multiple probes	14 (11)
Combined SCED	9 (7)
Others (e.g., AATD, withdrawal, random, between subjects vs. within subject)	9 (7)
Changing criterion	5 (4)

Note. Thirty-six Meta-analyses included more than one type of study designs; seven studies just included one type of study design.

of intervention effectiveness are included as moderator in 15 and 13 SCED meta-analyses respectively. Eleven SCED meta-analyses discuss instructional arrangement (i.e., how the instruction has been provided, such as individual to individual or individual to classroom); 11 incorporate publication type (e.g., journal articles, thesis or dissertation), and 10 SCED meta-analyses discuss validity. Furthermore, six SCED meta-analyses discuss the context of the setting (e.g., simulated setting or natural setting) and specifics about functional behavior assessment (FBA). Finally, five SCED meta-analyses discuss the effect of the study and the study findings as moderators.

*Measurement scale.* The measurement scale of these 13 study level moderators is nominal (See Table 2). Two moderators are coded inconsistently, namely the interobserver agreement and validity. Specifically, 11 SCED meta-analyses code the interobserver agreement as a nominal moderator, whereas four code this as continuous. Validity is coded as a nominal variable in nine SCED meta-analyses whereas one study considers it as continuous. In that study, validity is reported as the specific degree of validity.

*Missing data.* Most SCED meta-analyses report no missing data for moderators at the study level (i.e., study design, physical setting of intervention, design standards, design strength, interobserver agreement, publication type, instructional arrangement, improvement/findings, context, and FBA method) (see Table 2). For example, among the 42 SCED meta-analyses discussing study design, 29 report no missing data, while two SCED meta-analyses have

missing data and 11 SCED meta-analyses do not mention whether there is missing data. Among 38 SCED meta-analyses discussing physical settings of intervention, 28 SCED meta-analyses report no missing data, one SCED meta-analysis has missing data, and nine SCED meta-analyses do not mention whether there was missing data. More than half of the SCED meta-analysis including the moderators maintenance, generalization, and validity report missing data. For instance, ten SCED meta-analyses report missing data when discussing maintenance of intervention effects, while only two SCED meta-analyses report no missing data, and three SCED meta-analyses do not mention whether there was missing data or not.

### *Participant specific moderators*

*Ratio and type of moderators.* The average ratio between the number of participant level moderators and the number of units at the participant level is 0.06. This means that, on average, meta-analyses with 10 SCED study participants include less than one moderator. This ratio ranges from 0 to 0.24. A total of 18 unique study level moderators (across all SCED meta-analyses eligible for inclusion) are identified among which six are reported by at least five SCED meta-analytic studies. These six moderators represent commonly reported participant level moderators and characteristics of these are summarized in Table 2. An overview of the other, less commonly reported study level moderators, can be found in Supplemental Appendix D. The most popular participant moderator is age (55 out of the 60 SCED meta-analyses), followed by disability status ( $n=44$ ) and gender ( $n=39$ ). Seventeen SCED meta-analyses discuss ethnicity as a moderator and 17 studies report the participant's functional repertoires (e.g., delusional speech, hallucinatory speech, disorganized speech, different level of communication function, and different level of academic achievement) as a moderator. Finally, five SCED meta-analyses include whether participants received special education or not.

*Measurement scale.* Among the six participant specific moderators, four moderators (i.e., disability status, gender, ethnicity, and received special education) are included as nominal scaled variables. There is no consistency in the way the other two commonly encountered moderators are coded. Twenty-five SCED meta-analyses code age as a nominal variable, eight SCED meta-analyses consider it as ordinal and 25 SCED meta-analyses code it as continuous. Interestingly enough, among 55 SCED meta-analyses discussing age, three studies code age as both nominal and continuous. For functional repertoires, 14 out of 17 SCED meta-analyses code it as nominal, while three SCED meta-analyses code it as continuous (i.e., the degree of functional repertoire or relevant functional test scores).

*Missing data.* The majority of SCED meta-analyses report no missing data for three participant level moderators, namely age, disability status, and gender. For example, 30 SCED meta-analyses report no missing data for the moderator age, whereas 12 studies report missing data. The remaining 14 SCED meta-analyses do not mention any information about missing data for age. For disability status, 22 SCED meta-analyses report no missing data, nine SCED meta-analyses report having missing data, and 13 studies do not include information about missing data.

For the moderators ethnicity and functional repertoires, most SCED meta-analyses report missing data ( $n=13$  for ethnicity and  $n=8$  for functional repertoires). Only two SCED meta-analyses report no missing data for ethnicity, while this number equals six for functional repertoires. The number of SCED meta-analyses that do not provide information related to missing data equals two and three for ethnicity and functional repertoires respectively.

For the variable received special education, an equal amount of studies provide information related to missing data as do not provide information ( $n=2$ ). One SCED meta-analysis reports this moderator with missing data.

#### *Within-participant moderators: Intervention and outcome*

*Ratio and type of moderators.* The average ratio between the number of intervention moderators and the number of units at the within-participant level is 0.14. This means that, on average, meta-analyses with ten observations include at least one intervention specific moderator. This ratio ranges from 0 to 0.60. A total of 18 unique intervention specific moderators are identified among which six moderators are reported by at least five meta-analytic studies. Characteristics of these six moderators are presented in Table 2. Forty-four out of the sixty SCED meta-analyses include the intervention program as a moderator (e.g., video modeling, visual cueing, augmentative, or alternative communication). Another 28 SCED meta-analyses mention intervention agents as a moderator. This moderator indicates whether the agent delivering the intervention is a professional (i.e., researcher, clinician, or therapist), a classroom staff member, a student, or a parent. Twenty-one SCED meta-analyses include intervention technique as a moderator (e.g., reinforcement of appropriate behaviors or extinction of the problem behaviors), and 18 SCED meta-analyses discuss the intervention dosage (i.e., length and/or magnitude of intervention). Fourteen SCED meta-analyses include intervention fidelity and seven SCED meta-analyses mention the technology device used for intervention delivery or data collection.

The average ratio of the number of outcome moderators and the number of units at the within-participant level is 0.07. This means that, on average, meta-analyses with ten observations include no outcome specific moderator.

This ratio ranges from 0 to 0.60. A total of six unique outcome specific moderators are identified among which only two moderators are reported by at least five SCED meta-analyses. Characteristics of these two moderators are presented in Table 2. The most popular outcome specific moderator is outcome domain (i.e., 54 out of the 60 SCED meta-analyses). The moderator outcome domain refers to the specific outcome or the domain of the outcome that was measured in the primary studies, such as academic skills, adaptive skills, and emotion recognition. Fourteen out of the 60 SCED meta-analyses include the methods of measuring outcomes as moderator. This refers to the method, tool, or technique that was used to measure outcomes, such as student rating, teacher rating, or systematic direct observation.

*Measurement scale.* The measurement scale of all six most commonly included intervention specific moderators is nominal. However, two of these six moderators are coded inconsistently across SCED meta-analyses. Specifically, ten SCED meta-analyses code intervention dosage as a nominal variable, two as ordinal and six as continuous (e.g., the time of intervention). For fidelity, 13 SCED meta-analyses code it as nominal, while just one SCED meta-analysis code it as continuous (i.e., degree of intervention fidelity reflected as a percentage). The two commonly used outcome specific moderators (i.e., outcome domain and methods of measuring outcomes) are consistently coded in SCED meta-analyses as nominal.

*Missing data.* The majority of SCED meta-analyses report no missing data for six out of the eight within-participant moderators (i.e., intervention program, intervention agent, intervention techniques, technology devices, outcome domain, and methods of measuring outcomes). For instance, 26 SCED meta-analyses report no missing data for the intervention moderator, whereas this equals seven for having missing data and 11 for not mentioning whether missing data is present. For intervention agents, 13 SCED meta-analyses report no missing data, eight report having missing data and seven do not mention whether or not there was missing data. For the outcome domain, 35 SCED meta-analyses report no missing data, while five SCED meta-analyses report having missing data and 14 SCED meta-analyses do not mention the information about missing data

Most SCED meta-analyses do not provide information related to missing data for intervention dosage ( $n=9$ ). For dosage, five SCED meta-analyses report no missing data and the other four SCED meta-analyses report having missing data. For fidelity, most SCED meta-analyses report having missing data ( $n=7$ ), while three SCED meta-analyses report no missing data and four do not mention any information about missing data.

## Moderator Analysis

*Number of SCED meta-analyses quantifying moderation effects.* Among the 42 SCED meta-analyses discussing study design as moderator, about half ( $n=20$ ) report a metric to quantify its effect. A metric reflecting the effect of the moderator physical settings of intervention delivery is included in only 15 out of the 38 SCED meta-analyses. Less than one third (7 out of 23) of the meta-analyses discussing design standards report a metric to evaluate the effect of the moderator. Six out of the 11 SCED meta-analyses including instructional arrangement as a moderator report a metric. No more than five SCED meta-analyses report metrics for the other commonly encountered study level moderators (e.g., design strength, interobserver agreement, maintenance, and generalization).

For participant specific moderators, 34 out of 55 SCED meta-analyses that discuss age as a moderator report a metric quantifying the effect of that moderator. A metric is included for half of the SCED meta-analyses (22 out of 44) that focus on disability status. Nine SCED meta-analyses report a metric for gender, and the same number of SCED meta-analyses report a metric for functional repertoires. Only two out of the 17 SCED meta-analyses discussing ethnicity report a metric. For receiving special education, two out of five report a metric.

Regarding the intervention specific moderators, among 44 SCED meta-analyses discussing intervention program as a moderator, 34 studies report the metric for this moderator. Fifteen out of 28 SCED meta-analyses including intervention agents as a moderator report the metric to evaluate the effect of this moderator. Twelve out of 21 SCED meta-analyses discussing intervention techniques report the metric for intervention techniques. Among 18 SCED meta-analyses mentioning intervention dosage, nine report the metric for this moderator. Moreover, two out of 14 SCED meta-analyses discussing fidelity report the metric for fidelity. Two out of seven SCED meta-analyses report the metric for the moderator of technology devices.

In terms of outcome specific moderators, 41 out of 54 SCED meta-analyses discussing outcome domain report the metric for this moderator. Among 14 SCED meta-analyses analyzing methods of measuring outcomes as a moderator, three report the metric for methods of measuring outcomes.

*Metrics used to quantify moderator effects.* In total, 13 different types of metrics are reported in the reviewed SCED meta-analyses, which can be clustered into four broader categories, namely non-overlap metrics, regression based/ HLM metrics, the log ratio, and the standardized mean difference metrics. In order to enhance the discussion, these broader categories are summarized, but a complete overview of the specific metrics within each



of these four broader categories can be obtained upon request. This section presents the results for moderators included by the largest number of SCED meta-analysis per level (see Table 4) as similar findings apply for other moderators. Again, the complete overview can be obtained upon request by the first author. For within-participant level, intervention specific moderators and outcome specific moderators are reported separately. As a consequence, a total of four moderators are selected for in depth discussion. These moderators are study design (i.e., study level), age (i.e., participant level), intervention program (i.e., within-participant level–intervention specific) and domain of outcome (i.e., within-participant–outcome specific). The results show that 71% of SCED meta-analyses use non-overlap metrics, while 22% of studies use the SMD metrics. Only 4% use regression based metrics and 3% log ratio metrics. In sum, non-overlap metrics are most commonly used in SCED meta-analyses, followed by the SMD metric. Regression based metrics and log ratio metrics are less popular.

*Analysis approach used to summarize moderator effects across studies.* The type of analyses used in SCED meta-analyses to summarize moderator effects across studies, can be clustered within four broader categories; description in words, inclusion of frequency or frequency table, reporting of quantitative metric (e.g., reporting metric and 95% confidence interval), and inclusion of a statistical analysis (i.e., reporting statistical significance). The analysis approach for each moderator is presented in Table 2. The most commonly used analysis approach is description in words, followed by frequencies. The category of statistical analysis is the least used approach.

For study level moderators, at least 50% of studies analyze study design, publication type, interventional arrangement, context, and FBA method using a quantitative metric. A range of 20% to 46% of meta-analyses analyze physical settings of intervention, design standard, design strength, maintenance, generalization, validity, and improvement using a quantitative metric. For the moderator interobserver agreement, only 7% of studies use a quantitative metric. However, no studies analyzing interobserver agreement, generalization, validity, and improvement use a statistical analysis. Except from these four moderators, the percent of studies using statistical analysis to analyze other study level moderators ranges from 12% to 45%.

For participant level moderators, more than 50% of studies analyze age, disability status, functional repertoires, and received special education using a quantitative metric. Thirty-three percent of studies analyze gender using a quantitative metric, while 18% of studies analyze ethnicity using a quantitative metric. The percent of SCED meta-analyses studies using statistical analysis to analyze participant level moderators ranged from 6% to 60%.

**Table 4.** Number (%) of Meta-Analyses Using Specific Effects Sizes (The Most Popular Moderator in Each Moderator Level).

Four broad categories of metrics	Specific effect size	Study design	Intervention	Age	Outcome domain	Total number (%)
Non-overlap	Tau U	12	21	23	24	80 (41)
	PND/POD	5	4	4	5	18 (9)
	R-IRD/IRD	1	3	2	4	10 (5)
	NAP	2	3	3	2	10 (5)
	Phi coefficient	0	3	3	3	9 (5)
	PAND	0	2	2	3	7 (4)
	PEM	1	0	0	2	3 (1.5)
	Rho (intraclass correlation)	0	1	1	1	3 (1.5)
	PNCDD	1	0	0	1	2 (1)
	Subtotal					142 (73)
SMD	Hedges' g	5	7	5	4	21 (11)
	BC-SMD/SMD	4	7	5	4	20 (10)
Subtotal					41 (21)	
Regression based	Regression based effect size	1	2	2	2	7 (4)
Subtotal						7 (4)
Log ratio	Log response ratio (LRR)/response rate	1	3	0	1	5 (2)
Subtotal						5 (2)

For intervention specific moderators, at least 50% of studies analyze intervention program, intervention agent, intervention techniques, intervention dosage, and technology devices using a quantitative metric, while 14% of studies analyze fidelity using a quantitative metric. A range of 14% to 50% of studies analyze intervention specific moderators using statistical analysis. For outcome specific moderators, 78% of studies analyze outcome domain using a quantitative metric, while 59% of studies analyze this moderator using statistical analysis. Thirty-six percent of studies analyze outcome measurement using quantitative metric, and 21% of studies analyze it using statistical analysis.

## **Discussion**

### *Meta-Analysis of SCEDs*

Meta-analysis is a powerful technique for the quantitative synthesis of primary study results (Borenstein et al., 2009; Card, 2016; Cooper, 2017; Glass, 1976; Hedges & Olkin, 1985; Lipsey & Wilson, 2001). As research production is growing exponentially in the field of SCEDs, researchers, practitioners and policy makers are unable to read all research. Therefore, meta-analysis is a welcomed technique in the field of SCEDs. By combining research evidence across SCED studies (investigating the same intervention and the same outcome variable) using meta-analytic techniques, an objective summary statistic evaluating the effectiveness of an intervention can be obtained. Meta-analysis can be used to make more general conclusions related to the effectiveness of an intervention, reduces sampling error, and contributes to evidence-based decisions in practice, policy and research.

One meta-analytic that has been recommended for the quantitative summary of SCEDs is hierarchical linear modeling. The hierarchical linear model takes the multilayered data structure into account as SCED data (level 1) is nested within participants (level 2) and participants are nested within studies (level 3). By explicitly modeling these three levels, the source of systematic variability in intervention effectiveness between and within SCED studies can be identified and moderators at these different levels can be added in an effort to explain this systematic variability. This systematic review focuses on summarizing SCED meta-analyses including moderators as this is currently missing in the literature.

### *SCED Meta-Analysis including Moderators*

There is a lack of methodological research evaluating statistical properties of the hierarchical linear modeling approach summarizing SCED studies, with

the inclusion of moderators. The goal of this study is to provide an overview and description of commonly encountered moderator characteristics and analysis techniques that can be used to inform future methodological research as such an overview is currently missing. This systematic review of SCED meta-analysis provides a comprehensive overview and discussion of moderator characteristics typically included in SCED meta-analyses. Second, moderator analysis techniques are summarized (i.e., the metric used to quantify moderator effectiveness, the unit of analysis, and the specific approach used to combine metrics across cases and/or across studies).

### *General Moderator Characteristics*

This systematic review presents moderator characteristics based on a total of 60 SCED meta-analyses published between 2016 and 2019. Based on the systematic review, the following moderator characteristics can be considered in future methodological work. At the study level, the most popular moderators are study design, physical setting of intervention, design standards, design strength, interobserver agreement, and maintenance. At the participant level, commonly encountered moderators are age, disability status, gender, ethnicity, and functional repertoires. Finally, at the within-participant level, the most discussed moderators are outcome domain, intervention program, intervention agent, intervention techniques, and intervention dosage. All of the aforementioned moderators' measurement scales are nominal, except from interobserver agreement, age, functional repertoires and intervention dosage. Almost all SCED meta-analyses reported no missing data for included moderators. This implies that a condition representing missing data for moderators does not necessarily need to be part of future methodological work as this is not an issue.

### *Number of Moderators per Units*

Current systematic review provides insights into the typical ratio of the number of moderators versus the number of units at the study level, participant level, and within-participant level. If a SCED meta-analysis includes ten primary SCED studies and a total of ten participants among these ten primary SCED studies, typically this SCED meta-analysis discusses two study level moderators, no more than one participant level moderator (i.e., ratio = .06), and one moderator at the within-participant level. The ratio of the number of moderators at the study level to the number of units at that level ranges from 0 to 1.4. This indicates that some SCED meta-analyses do not include moderators at the study level, whereas others include more moderators relative to

the number of units at that level. It is recommended that future simulation studies include conditions reflecting both extremes. The ratio of the number of moderators at the participant level to the number of units at that level ranges from 0 to 0.24. This indicates that an SCED meta-analysis including on average 10 participants per study, typically includes up to two moderators. Only one meta-analysis did not include any participant level moderators. The current study finds that the ratio at the within-participant level ranges from 0 to 0.60 (both for intervention specific and outcome specific moderators).

### *Quantification of Moderators*

Although previous research found that most meta-analyses conducted a moderator analysis (Jamshidi et al., 2020), details of the analysis approach related to each moderator per level was not reported. Current study indicates that statistical significance testing is the least commonly used approach to analyze moderators. The majority of SCED meta-analyses used quantitative metrics to analyze moderator effects. The most commonly used metric to quantify moderator effectiveness is Tau-U. This metric is considered to be a more advanced non-overlap statistic as it compares all the baseline observations with all the intervention observations and it has the potential to account for baseline trends (Parker, Vannest, Davis, & Sauber, 2011). However, recent methodological research has indicated that this metric has no meaningful scale (i.e., is not bounded between -1 and 1), is biased in certain conditions, and is difficult to interpret (Fingerhut et al., 2021; Tarlow, 2017). In addition, synthesizing moderator effects using the average Tau-U ignores the nested data structure and is not recommended. The hierarchical linear modeling approach deals with all these issues, but is only used in a limited number of SCED meta-analyses. As a consequence, efforts are needed to further disseminate the approach to a broader audience.

### *Limitations and Future Research*

The review of previously published systematic reviews of SCED meta-analyses revealed that items related to moderators have not been considered for inclusion in reporting guidelines to assess the quality of SCED meta-analyses and systematic reviews. Therefore, future research is needed to develop such items. Reporting guidelines and checklists for quality assessment of group-comparison design meta-analyses and primary level SCEDs embed to some extent items related to moderator characteristics. Therefore, these items can be considered as a starting point. Having access to information related to moderator characteristics can have far reaching implications for practice,

policy and theory as some interventions might only be effective given a certain set of study and/or participant characteristics.

Future Monte-Carlo simulation studies are needed to provide recommendations about the number of units at the different levels of the hierarchical liner modeling approach that are needed to identify true intervention and moderator effects, given a certain set of design conditions representative for SCED meta-analysis including moderators. The moderator characteristics reported in current systematic review can inform these design conditions. However, parameter values for moderator effects (e.g., the size of the effect of age on the intervention), correlations between the moderators (e.g., correlation between age and gender) and correlations between the moderators and the intervention (e.g., correlation between age and intervention) are not discussed. Identifying these parameter values is recommended for future research.

This systematic review presents moderator characteristics based on a total of 60 SCED meta-analyses published between 2016 and 2019. SCED meta-analyses prior to 2016 could also been explored. However, current systematic review evaluates 910 SCED meta-analysis and provides insights in commonly used SCED moderators based on the most recent SCED meta-analyses.

### *Implications of SCED Meta-Analysis for Evidence-based Practice*

In sum, this systematic review provides a comprehensive overview and discussion of moderator characteristics typically included in SCED meta-analyses. This overview of moderator characteristics is timely and can inform the design conditions to be included in future methodological work. Future methodological work is needed to provide answers to practical questions when designing SCED meta-analyses such as: (a) What are typically encountered study, participants and within-participants moderators? (b) What is the scale of typically encountered moderators? (c) What is the number of moderators relative to the number of units? and (d) What is the anticipated power to detect true intervention and moderator effects given a set of design conditions representative for SCED meta-analyses? By further enhancing the field of SCED meta-analysis by including moderators, evidence can be obtained about what intervention works, when, where, for who and at which cost.

### **Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Institute of Education Sciences, U.S. Department of Education, through grant R305D190022. The content is solely the responsibility of the author and does not necessarily represent the official views of the Institute of Education Sciences, or the U.S. Department of Education.

## ORCID iD

Mariola Moeyaert  <https://orcid.org/0000-0003-1453-8162>

## Supplemental Material

Supplemental material for this article is available online.

## References

- Baek, E. K., & Ferron, J. M. (2013). Multilevel models for multiple-baseline data: Modeling across-participant variation in autocorrelation and residual variance. *Behavior Research Methods*, *45*(1), 65–74. <https://doi.org/10.3758/s13428-012-0231-z>
- Beretvas, S. N., & Chung, H. (2008). A review of meta-analyses of single-subject experimental designs: Methodological issues and practice. *Evidence-Based Communication Assessment and Intervention*, *2*(3), 129–141.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons.
- Card, N. A. (2016). *Applied meta-analysis for social science research*. Guilford.
- Cooper, H. (2017). *Research synthesis and meta-analysis: A step-by-step approach*. SAGE Publications, Inc.
- Council for Exceptional Children (CEC). (2014). *Council for exceptional children standards for evidence-based practices in special education*. Council for Exceptional Children. <https://journals.sagepub.com/doi/10.1177/0040059914531389>
- Declercq, L., Jamshidi, L., Fernández-Castilla, B., Beretvas, S., Moeyaert, M., Ferron, J., & Van den Noortgate, W. (2019). Analysis of single-case experimental count data using the linear mixed effects model: A simulation study. *Behavior Research Methods*, *51*(6), 2477–2497. <https://doi.org/10.3758/s13428-018-1091-y>
- Farmer, J. L., Owens, C. M., Ferron, J. M., & Allsopp, D. H. (2010). *A methodological review of single-case meta-analyses* [Paper presentation]. American Educational Research Association, Denver, CO, United States.
- Ferron, J., & Scott, H. (2005). Multiple baseline designs. In B. Everitt & D. Howell (Eds.), *Encyclopedia of behavioral statistics* (Vol. 3, pp. 1306–1309). Wiley & Sons Ltd.
- Ferron, J. M., Bell, B. A., Hess, M. F., Rendina-Gobioff, G., & Hibbard, S. T. (2009). Making treatment effect inferences from multiple-baseline data: The utility of

- multilevel modeling approaches. *Behavior Research Methods*, 41(2), 372–384. <https://doi.org/10.3758/BRM.41.2.372>
- Fingerhut, J., \*Xinyun, X., & Moeyaert, M., (2021). Impact of within-case variability on Tau-U and regression-based effect size measures for single-case experimental data. *Evidence-Based Communication and Intervention*.
- Glass, G. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(10), 3–8. Retrieved March 8, 2021, from <http://www.jstor.org/stable/1174772>
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press.
- Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2012). A standardized mean difference effect size for single case designs. *Research Synthesis Methods*, 3(3), 224–239.
- 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.
- Heyvaert, M., Maes, B., Van den Noortgate, W., Kuppens, S., & Onghena, P. (2012). A multilevel meta-analysis of single-case and small-n research on interventions for reducing challenging behavior in persons with intellectual disabilities. *Research in Developmental Disabilities*, 33(2), 766–780.
- Heyvaert, M., Saenen, L., Maes, B., & Onghena, P. (2014). Systematic review of restraint interventions for challenging behaviour among persons with intellectual disabilities: Focus on effectiveness in single-case experiments. *Journal of Applied Research in Intellectual Disabilities*, 27(6), 493–510.
- Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, 71(2), 165–179.
- Hurwitz, J. T., Kratochwill, T. R., & Serlin, R. C. (2015). Size and consistency of problem-solving consultation outcomes: An empirical analysis. *Journal of School Psychology*, 53(2), 161–178.
- Jamshidi, L., Heyvaert, M., Declercq, L., Fernández-Castilla, B., Ferron, J., Moeyaert, M., Beretvas, S. N., & Van den Noortgate, W. (2018). Methodological quality of meta-analyses of single-subject experimental studies. *Research in Developmental Disabilities*, 79, 97–115.
- Jamshidi, L., Heyvaert, M., Declercq, L., Fernández-Castilla, B., Ferron, J. M., Moeyaert, M., & Van den Noortgate, W. (2020). A systematic review of single-case experimental design meta-analyses: Characteristics of study designs, data and analyses. *Evidence-Based Communication Assessment and Intervention*.
- Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings*. Oxford University Press.
- Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single-case designs technical documentation. Retrieved from What Works Clearinghouse website [http://ies.ed.gov/ncee/wwc/pdf/wwc\\_scd.pdf](http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf)



- Kratochwill, T. R., & Levin, J. R. (2010). Enhancing the scientific credibility of single-case intervention research: Randomization to the rescue. *Psychological Methods, 15*(2), 124–144. <https://doi.org/10.1037/a0017736>
- Kung, J., Chiappelli, F., Cajulis, O. O., Avezova, R., Kossan, G., Chew, L., & Maida, C. A. (2010). From systematic reviews to clinical recommendations for evidence-based health care: Validation of revised assessment of multiple systematic reviews (R-AMSTAR) for grading of clinical relevance. *The Open Dentistry Journal, 4*, 84–91.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE Publications, Inc.
- Lobo, M. A., Moeyaert, M., Cunha, A. B., & Babik, I. (2017). Single-case design, analysis, and quality assessment for intervention research. *Journal of Neurologic Physical Therapy: JNPT, 41*(3), 187.
- Logan, L. R., Hickman, R. R., Harris, S. R., & Heriza, C. B. (2008). Single-subject research design: Recommendations for levels of evidence and quality rating. *Developmental Medicine & Child Neurology, 50*(2), 99–103.
- Maggin, D. M., O’Keeffe, B. V., & Johnson, A. H. (2011). A quantitative synthesis of methodology in the meta-analysis of single-subject research for students with disabilities: 1985–2009. *Exceptionality: A Special Education Journal, 19*(2), 109–135.
- Moeyaert, M., Ferron, J. M., Beretvas, S. N., & Van den Noortgate, W. (2014). From a single-level analysis to a multilevel analysis of single-case experimental designs. *Journal of School Psychology, 52*(2), 191–211. <https://doi.org/10.1016/j.jsp.2013.11.00>
- Moeyaert, M., Ugille, M., Ferron, J. M., Beretvas, S. N., & Van den Noortgate, W. (2013a). Modeling external events in the three-level analysis of multiple-baseline across-participants designs: A simulation study. *Behavior Research Methods, 45*(2), 547–559. <https://doi.org/10.3758/s13428-012-0274-1>
- Moeyaert, M., Ugille, M., Ferron, J. M., Beretvas, S. N., & Van den Noortgate, W. (2013b). The three-level synthesis of standardized single-subject experimental data: A Monte Carlo simulation study. *Multivariate Behavioral Research, 48*(5), 719–748.
- Moeyaert, M., Ugille, M., Ferron, J. M., Beretvas, S. N., & Van den Noortgate, W. (2014). Three-level analysis of single-case experimental data: Empirical validation. *The Journal of Experimental Education, 82*(1), 1–21.
- Moeyaert, M., Ugille, M., Ferron, J. M., Beretvas, S. N., & Van den Noortgate, W. (2016). The misspecification of the covariance structures in multilevel models for single-case data. *Journal of Experimental Education, 84*(3), 473–509.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), e1000097.
- Parker, R. I., & Vannest, K. J. (2009). An improved effect size for single case research: Nonoverlap of all pairs (NAP). *Behavior Therapy, 40*(4), 357–367. <https://doi.org/10.1016/j.beth.2008.10.006>

- Parker, R. I., Vannest, K. J., & Brown, L. (2009). The improvement rate difference for single-case research. *Exceptional Children, 75*(2), 135–150. <https://doi.org/10.1177/001440290907500201>
- Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Nine non-overlap techniques for single case research. *Behavior Modification, 35*(4), 303–322. <https://doi.org/10.1177/0145445511399147>
- 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. <https://doi.org/10.1016/j.beth.2010.08.006>
- Petit-Bois, M., Baek, E. K., Van den Noortgate, W., Beretvas, S. N., & Ferron, J. M. (2016). The consequences of modeling autocorrelation when synthesizing single-case studies using a three level model. *Behavior Research Methods, 48*(2), 803–812.
- Pustejovsky, J. E. (2015). Measurement-comparable effect sizes for single-case studies of free-operant behavior. *Psychological Methods, 20*(3), 342–359. <https://doi.org/10.1037/met0000019>
- Reichow, B., Volkmar, F. R., & Cicchetti, D. V. (2008). Development of the evaluative method for evaluating and determining evidence-based practices in autism. *Journal of Autism and Developmental Disorders, 38*(7), 1311–1319.
- SAS Institute Inc. (2013). *SAS® 9.4 Statements: Reference*. SAS Institute Inc.
- Schlosser, R. W., Lee, D. L., & Wendt, O. (2008). Application of the percentage of non-overlapping data (PND) in systematic reviews and meta-analyses: A systematic review of reporting characteristics. *Evidence-Based Communication Assessment & Intervention, 2*(3), 163–187.
- Schlosser, R. W., Sigafos, J., & Belfiore, P. (2009). EVIDAAC comparative single-subject experimental design scale (CSSEDARS).
- Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single subject research: Methodology and validation. *Remedial and Special Education, 8*(2), 24–33.
- 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, 18*(3), 385–405.
- Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., Porter, A. C., Tugwell, P., Moher, D., & Bouter, L. M. (2007). Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology, 7*(1), 10.
- Stone, B. A. (2011). *The effect of physical activity on youths' cognitive, academic, and behavioral outcomes: A meta-analysis of single case design studies* [Doctoral dissertation]. ProQuest Dissertations and Theses. (Accession Order No. 1791982098).
- Tarlow, K. R. (2017). An improved rank correlation effect size statistic for single-case designs: Baseline corrected Tau. *Behavior Modification, 41*(4), 427–467. <https://doi.org/10.1177/0145445516676750>
- Tate, R. L., McDonald, S., Perdices, M., Togher, L., Schultz, R., & Savage, S. (2008). Rating the methodological quality of single-subject designs and n-of-1

- trials: Introducing the Single-Case Experimental Design (SCED) Scale. *Neuropsychological Rehabilitation*, 18(4), 385–401.
- Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Barlow, D. H., Horner, R., Kazdin, A., Kratochwill, T., McDonald, S., & Sampson, M. (2016). The single-case reporting guideline in behavioural interventions (SCRIBE) 2016 statement. *Remedial and Special Education*, 37(6), 370–380.
- Ugille, M., Moeyaert, M., Beretvas, S. N., Ferron, J., & Van den Noortgate, W. (2012). Multilevel meta-analysis of single-subject experimental designs: A simulation study. *Behavior Research Methods*, 44(4), 1244–1254. <https://doi.org/10.3758/s13428-012-0213-1>
- Van den Noortgate, W., & Onghena, P. (2003a). Hierarchical linear models for the quantitative integration of effect sizes in single case research. *Behavior Research Methods, Instruments & Computers*, 35(1), 1–10. <https://doi.org/10.3758/BF03195492>
- Van den Noortgate, W., & Onghena, P. (2003b). Multilevel meta-analysis: A comparison with traditional meta-analytical procedures. *Educational and Psychological Measurement*, 63(5), 765–790. <https://doi.org/10.1177/0013164403251027>
- 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), 142–158. <https://doi.org/10.1080/17489530802505362>
- Vanderkerken, L., Heyvaert, M., Maes, B., & Onghena, P. (2013). Psychosocial interventions for reducing vocal challenging behavior in persons with autistic disorder: A multilevel meta-analysis of single-case experiments. *Research in Developmental Disabilities*, 34, 4515–4533.
- Wang, S. Y., Parrila, R., & Cui, Y. (2013). Meta-analysis of social skills interventions of single-case research for individuals with autism spectrum disorders: Results from three-level HLM. *Journal of Autism and Developmental Disorders*, 43(7), 1701–1716.
- What Works Clearinghouse (WWC). (2020). *What works clearinghouse standards handbook* (Version 4.1). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. This report is available on the What Works Clearinghouse website at <https://ies.ed.gov/ncee/wwc/handbooks>
- Zimmerman, K. N., Pustejovsky, J. E., Ledford, J. R., Barton, E. E., Severini, K. E., & Lloyd, B. P. (2018). Single-case synthesis tools II: Comparing quantitative outcome measures. *Research in Developmental Disabilities*, 79, 65–76. <https://doi.org/10.1016/j.ridd.2018.02.001>

## Author Biographies

**Mariola Moeyaert** is an associate professor in Educational Psychology and Methodology at the University at Albany. Dr. Mariola Moeyaert obtained her PhD in Educational Sciences from the University of Leuven in 2014. Major research interests and publications are in the field of multilevel analysis, meta-analysis, and interrupted time series analysis. She has already (co)authored more than 50 international

publications, for instance in *Psychological Methods*, *Multivariate Behavior Research* and *Behavior Research Methods*. She is currently conducting research as PI on an early career grant awarded by the Institute of Education Sciences (“Assessing Generalizability and Variability of Single-Case Design Effect Sizes Using Multilevel Modeling Including Moderators”).

**Panpan Yang** is a PhD candidate in the Educational Psychology and Methodology program at the University at Albany. She got her master degree from that same program. Her major research interests are in the domain of adolescent and preschoolers’ development and parenting, and advanced methodologies (e.g., multilevel modeling and growth modeling).

**Xinyun Xu** is a fourth year PhD student in the Educational Psychology and Methodology at the University at Albany. She holds an M.S. in Educational Psychology and Methodology from the Department of Educational and Counseling Psychology at the University at Albany. Her primary research interest is in single-case experimental design, multilevel-analysis, meta-analysis, and non-parametric statistics.

**Esther Kim** is a fourth year PhD student in the Educational Psychology and Methodology at the University at Albany. She holds an M.S. in Educational Psychology and Methodology from the Department of Educational and Counseling Psychology at the University at Albany. She previously earned her bachelor’s degree in U.S. History. Her research interests include development and analysis of instruments, well-being of immigrants’ population, and self-regulated learning. She is currently working as research assistant at the Center for Women in Government and Civil Society (CWGCS).