

# What Works Clearinghouse

## Study Review Guide Instructions for Reviewing Single-Case Designs

Released March 2021

This document provides step-by-step instructions on how to complete the Study Review Guide (SRG, S4.1 V2; March 2021, release) for single-case designs (SCDs).

For every What Works Clearinghouse (WWC) review, reviewers will be asked to complete an SRG. A completed SRG should be a reviewer's independent assessment of the study, relative to the criteria specified in the review protocol and the *WWC Standards and Procedures Handbooks*. At the end of the review process, a Master Review Guide (MRG) will represent the final assessment of the study and provide a record for the WWC, as well as serve as a key input in producing reports. For more details on the review process, see the instructions provided in the *Procedures Handbook* or the review protocol.

This guide is intended to be used by individuals trained and certified in the version 4.1 *WWC Standards and Procedures Handbooks* as well as version 4.1 SCD Standards. The current versions of the *WWC Standards and Procedures Handbooks* can be accessed at <https://ies.ed.gov/ncee/wwc/Handbooks>, review protocols can be accessed at <https://ies.ed.gov/ncee/wwc/Handbooks#protocol>, and the most current version of the SRG can be accessed at <https://ies.ed.gov/ncee/wwc/StudyReviewGuide>.

## Definition of a Study in SCD Research

The WWC characterizes all SCD experiments in the same publication as one study, and thus one study/publication can have multiple SCD experiments included within it. For example, a study could include three separate ABAB experiments (i.e., treatment reversal/withdrawal designs) for one student when there are data for three different eligible outcomes. Alternatively, a study can include three separate ABAB experiments for three separate eligible students when there are data for one eligible outcome. Each eligible SCD experiment should receive a WWC design rating (or disposition code) in the SRG. The study also will receive an overall study rating that the citation list will report.

## Navigating and Adding Rows to the SRG

There are seven worksheets (that is, tabs) in the SCD SRG: (1) SRG Information (2) *Main*, (3) *Review*, (4) *Data*, (5) *Plots*, (6) *Author Query & Response*, and (7) *Summary*. The SRG Information tab does not need to be filled by reviewers, it contains only SRG version information. Reviewers will need to add rows to the *Main Tab* to record information and disposition codes when there is more than one eligible SCD experiment within a study. Reviewers may also need to add rows to the *Review Tab* and/or *Author Query & Response Tab* when the number of recorded rows exceeds what is provided in the SCD SRG. Reviewers should be aware that the row numbers referenced throughout this user guide correspond to the original SCD SRG template. If reviewers add rows during the review process, the row numbers referenced in this guide will no longer directly align to your working version of the SCD SRG.

To add additional rows in any worksheet, start by clicking on the *Review Tab*, and select “Unprotect Sheet.” You can copy the rows you want to add by selecting them with your cursor, right-clicking, and selecting “Copy.” Then, place your cursor in the row below where you want to add them, right-click, and select “Insert Copied Cells.” Always use the “Copy” and “Insert Copied Cells” functions to preserve text, drop-down menus, and formatting in the SRG. Repeat the process to add as many rows as you need in the various tabs. When you are done adding rows, click “Protect Sheet” to ensure you do not overwrite a formula by mistake.

### ***General Characteristics of the SRG***

In the *Main Tab*, there are two places where a reviewer may need to add rows:

1. To represent all appropriate disposition codes for why the study is identified as *Not Eligible for Review* either for the study as a whole, or for a specific SCD experiment (row 16 and row 17).
2. To provide a study rating for each SCD experiment within a study (row 61 and row 62).

To add additional rows to capture the disposition, click on the *Review Tab*, and click “Unprotect Sheet.” Highlight rows 16 and 17, columns A through D, and select “Copy.” Right-click in the row below (that is, cell A18) and select “Insert copied cells.” If an “Insert Paste” dialog box pops up, select “Shift cells down” and click “OK.” You may need to adjust the row height of the inserted cells to view the cell contents. On the *Review Tab*, click “Protect Sheet” to ensure you do not overwrite a formula by mistake.

The *Review Tab* includes 20 rows for outcomes (rows 2–21). If the study you are reviewing has more than 20 outcomes × participants × time period, click on the *Review Tab*, and click “Unprotect Sheet.” Highlight a blank row. Press Ctrl + C. Right-click and select “Insert Copied Cells.” Repeat to add as many rows as you need. On the *Review Tab*, click “Protect Sheet” to ensure you do not overwrite a formula by mistake.

### ***Main Tab of the SRG***

The *Main Tab* of the SRG captures a prose summary and assessment of the study and is structured in three separate stages, described in detail below.

As a reviewer, one of your roles is to provide complete information for each element in a given section of the SRG. Reviewers are asked to enter data into columns B, C, and D in the *Main Tab*.

- Column B typically requires the reviewer to select the appropriate answer to a question posed in column A via a drop-down menu.
- Column C will require a written justification for the short answer in column B.
- Column D will require the reviewer to record the page numbers from the study to serve as the reference for the justification presented in column C.

#### ***Stage 1: Preliminary Screening***

Stage 1 of the study review assesses whether the study meets the WWC’s SCD screening criteria and whether the study is eligible for WWC review under a given review protocol. All WWC reviews are conducted using a specific review protocol. Due to differences between review protocols, screening and review decisions may differ depending on the protocol used in the review. For instance, a study may be eligible for review under one protocol, but ineligible under another due to differences in the sample or outcome measure requirements written in each protocol. Similarly, a study may meet standards under one protocol, but not meet standards for another, due to aspects of the research design that are considered differently under different review protocols.

Now, we will introduce each item housed in Stage 1 of the Study Review Guide, found in rows 2–23.

- **Full Citation** (cell A1)
  - Enter the full citation for the study.

Example:

Beard, K. Y., & Sugai, G. (2004). First step to success: An early intervention for elementary children at risk for antisocial behavior. *Behavioral Disorders*, 29(4), 396–409.

- **Study ID** (cell B1)
  - Enter the Study ID for the review.
- **Reviewer Number** (cell C1)
  - Enter your reviewer number. If an MRG, enter “MRG” into this cell.

- **Review Date** (cell D1)
  - Enter the date of your review.

### Overview

- **Standards and Protocol** (row 5)

Enter the version of the WWC Standards being used for the review (cell B5). Enter the name and version of the protocol being used for the review (cell C5).

- **Intervention Name** (row 6)

Clearly indicate the name of the intervention(s) you are reviewing within this SRG (cell B6). Note whether a single name refers to multiple versions of an intervention or different names refer to the same or related interventions (cell C6).

Example:

(cell C6) *The study examined the effectiveness of Reading Recovery implemented as a pullout program for struggling readers. Students participated in 30-minute lessons four days a week.*

### Screening Criteria

- **Effectiveness.** Does the study examine the effect of an intervention? (row 8)

In cell B8, select “Yes” if the study claims to examine the effect of an intervention within the scope of the review, regardless of the quality of the design; select “No” otherwise.

- **Design.** Does the study use an SCD? (row 9)

SCDs have the following features: (a) an individual case serves as the unit of intervention administration and data analysis; (b) the case provides its own control for purposes of comparison; and (c) the outcome variable is measured repeatedly within and across different conditions or levels of the independent variable (that is, phases). In cell B9, select “Yes” if the study uses an SCD; select “No” if the study uses a design other than an SCD, such as a randomized controlled trial (RCT), a quasi-experimental design (QED), a regression discontinuity design (RDD), a pre-post design that lacks a comparison group, a meta-analysis, or a literature review. In column C, select the type of design; provide further details in Column D if appropriate. If the design is a comparison group design (RCT or QED) or an RDD, stop entering information using the SCD SRG template. Provided that you are certified to conduct a review for the given design, obtain access to the appropriate SRG for the RCT/QED or RDD, and complete the SRG in the correct format. If needed, please route the study in question to the leaders of the review so they can obtain a reviewer certified in the appropriate design standards.

- **Focus.** Is the intervention a program, product, policy, or practice whose primary focus is aligned with the review protocol? (row 10)

In cell B10, select “Yes” if the intervention meets the criteria for inclusion specified in the protocol under *Types of Interventions to Be Included*; select “No” otherwise.

In cell C10, briefly describe the intervention as it will be reviewed within the SRG.

- **Sample Alignment.** Does the study meet the requirements for sample characteristics specified in the review protocol? (row 11)

In cell B11, select “Yes” if the sample meets the criteria for inclusion specified in the protocol under *Types of Populations to Be Included* (for example, percent of English language learners or percent of general education); select “No” otherwise. In column C, describe the sample with respect to how it satisfies or does not satisfy this requirement. If the study is rated either *Meets WWC SCD Standards Without Reservations* or *Meets WWC SCD Standards With Reservations*, a more detailed description of the study participants will be provided in Stage 3 of the SRG.

Examples:

(Yes) *Each child had a diagnosis of autistic disorder, provided by an outside professional, according to the criteria of the American Psychiatric Association DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, 4th edition) (APA, 1994). Diagnoses were confirmed using the Autism Diagnostic Interview, Revised (ADI-R; Lord, Rutter, & Le Couteur, 1994) and the Autism Diagnostic Observation Schedule (ADCS; Lord, Rutter, DiLavore, & Risi, 1999).*

(Yes) *Donna was an eight-year-old girl with severe intellectual disability and cerebral palsy; Denise was a nine-year-old student with a severe intellectual disability who was legally blind. Ralph was a six-year-old student with a severe intellectual disability and cerebral palsy.*

- **Time.** Was the study published within the time frame established in the review protocol? (row 12)

In cell B12, select “Yes” if the study falls within the time frame outlined in the protocol; select “No” otherwise.

- **Age or Grade Range.** Does the study examine students in the age or grade range specified in the review protocol? (row 13)

In cell B13, select “Yes” if the sample meets the criteria for the age or grade range specified in the protocol under *Types of Populations to Be Included*; select “No” otherwise. In column C, please indicate (a) the grade levels of the students in the sample, or (b) the age range of the children in the sample, if the sample is described in terms of age rather than grade level. If the study receives a rating of either *Meets WWC SCD*

*Standards Without Reservations* or *Meets WWC SCD Standards With Reservations*, a more detailed description of the study participants will be provided in Stage 3 of the SRG.

Example:

*(Yes) Six participants were two to four years old (their ages ranged from 24 to 47 months).*

- **Location.** Does the study examine sample members in a location specified for the review protocol? (row 14)

In cell B14, select “Yes” if the study sample was drawn from the geographic region described in the protocol under *Types of Populations to Be Included*; select “No” otherwise. In cell C14, describe the location. If the study is rated either *Meets WWC SCD Standards Without Reservations* or *Meets WWC SCD Standards With Reservations*, a more detailed description of the study setting will be provided in Stage 3 of the SRG.

- **Outcomes.** Does the study include at least one eligible outcome in a review protocol domain? (row 15)

In cell B15, select “Yes” if the study estimates the impacts of the intervention on at least one outcome that falls into one of the eligible domains specified in the protocol under *Types of Outcomes to Be Included*. In column C (cell C15), list all outcomes included in the study, note which are eligible for review, and cite the relevant domain. You do not have to provide detailed information for each outcome here.

Example:

*(Yes) The study reported on two outcomes, percent time on task and prosocial behaviors.*

Select “No” if there are no outcomes that fall within the domains of the review protocol. In column C, briefly list available outcomes. Provide any additional documentation needed for a reviewer to confirm that the outcomes are not eligible under that protocol. If you are working on a WWC team and are uncertain whether any of the study’s outcomes qualify under the protocol, seek additional guidance from your review leadership.

Example:

*(No) The outcome is not eligible for review under the Students With Learning Disabilities review protocol. Although the study examined reading fluency, the outcome was measured in Spanish.*

- **Does the study meet WWC SCD eligibility requirements and review protocol screening criteria?** (row 16)

If more than one disposition is appropriate, copy and paste this row and select the additional disposition code.

From the drop-down menu (“Yes/No”) in column B, select “Yes” if the study met all of the screening criteria in this section and warrants a full review to determine whether the study design meets WWC SCD Standards. In column C, select “The study is eligible for review. An explanation for that decision is below” from the drop-down menu.

Select “No” in column B if the study failed to meet any of the screening criteria in this section, and select the appropriate screening disposition code in column C.

The screening disposition codes are as follows:

... is out of the scope of the protocol. This covers scenarios in which the study did not (a) use an intervention reported in English, (b) occur within the time frame specified in the protocol, (c) occur within a geographic area specified in the protocol, (d) occur within a setting specified in the protocol, or (e) include an outcome within a domain specified in the protocol.

... does not use a sample aligned with the protocol. This covers scenarios in which (a) the study did not use a sample within the age or grade range specified in the protocol, or (b) [the study did not show/the authors could not confirm/the WWC could not confirm] that at least 50 percent of the sample was classified [review specific classification].

... does not use an eligible design. This covers scenarios in which the study did not (a) examine the effectiveness of the intervention, (b) contain a primary analysis, (c) use an eligible SCD design, or (d) provide adequate or consistent information to assess whether it was eligible for review.

- **Explanation for Screening Disposition** (row 17)

If you selected “Yes” in cell B16, provide a brief explanation of why the study is eligible for review in cell C17.

Example:

*The study is eligible for review under the Intellectual Disabilities review protocol. The study is an SCD that examines the impacts of customized checklists on general housekeeping tasks for high school–age students diagnosed with intellectual disabilities.*

If you selected “No” in cell B16, provide a complete explanation for the screening disposition you selected in cell C17. One option is to consider using the scenarios described in the examples as the starting point for your explanation.

Example:

*The study is not eligible for review under the Children Identified With or at Risk for an Emotional Disturbance review protocol because the study participants were four years old and attended preschool at the time of the study.*



**Stage 2: Quality of Evidence**

If a study passes the preliminary screening in Stage 1, it is eligible to move on to Stage 2. Stage 2 of the review determines whether the study meets WWC standards. In this section, you will be making assessments about the level of evidence from the study. To improve the transparency of the review process, whenever possible, use direct quotes from the study when asserting that a study does not pass a particular standard.

*Graphical or Tabular Availability of Data (row 20)*

- **Is there a graphical illustration or a table of the raw data available for at least one experiment?** (row 21)

Select “Yes” or “No” in cell B21. In cell C21, describe any presentation of the data. If data are presented for some outcomes or contexts, but not for others, describe the outcomes or contexts for which the presentation of data are missing.

Examples:

(Yes) *The authors provided complete plots for all outcomes (problem behavior and on-task behavior).*

(Yes) *The authors provided plots of externalizing behaviors for Participants 1–4; however, they did not provide plots or tables for Participants 5–8.*

*Information About the Study Design(s) and Unit of Analysis (row 22)*

- **Does the study use a reversal/withdrawal, multiple baseline, multiple probe, alternating treatment, changing criterion, or other design?** Rows 23 through 28 ask for details regarding the SCD variations in the study. Some SCDs will only use one variation, while others might use more than one variation.

In column B for rows 23–27, select “Yes” if the corresponding design (reversal/withdrawal, multiple baseline, multiple probe, alternating treatment, or changing criterion) was used within the study and provide any important details about the study design(s) in column C. Otherwise, select “No.” If the study also contains a design not listed, select “Yes” in cell B28, and provide details regarding the design in cell C28.

In some cases, an SCD experiment could simultaneously use more than one type of design (such as a multiple baseline design that also could be treated as three separate ABAB designs). In this case, you should consult with the leaders of your review to decide which design WWC will use to characterize the findings. Select the design you have decided to use in this section and include a note about the decision.

- **Describe the unit of analysis.** (row 29)

Provide details regarding the unit of analysis in cell C29. Individual participants are the most common unit of analysis in an SCD, but the unit of analysis also might be a group

such as a dyad or classroom. If the unit of analysis is not an individual, please provide a description of the unit of analysis and how the samples were formed in cell C29.

Examples:

1. *The unit of analysis for this study was the individual participants: John, Mary, and Rebecca.*
2. *The unit of analysis for this study was made up of a dyad containing one student with a learning disability and one student helper, trained to assist the student to remain on task.*
3. *The unit of analysis was the three kindergarten classrooms engaged in the group contingency intervention.*

*Determine if Outcomes Meet WWC Requirements*

*Complete Outcome Details and Screening Section of Review Tab (columns A–M, row 31)*

Click on cell A31 to be taken to the beginning of the *Review Tab*. Complete the indicated section (instructions begin on page 3).

- **For all participants, did the authors collect interassessor agreement (IAA) for each eligible outcome at least once in each phase for each case?** (row 32)

Select “Yes” in cell B32 if IAA was assessed at least once for each phase within each outcome. Select “No” if it is clear that this requirement is not met. If it is not clear that this requirement is met, select “Uncertain” and note in cell C32 the need for an author query.

If “Uncertain” was selected in cell B32, note the reason for uncertainty in cell C32, and draft any author questions you have on the *Author Query & Response Tab*. If this IAA information is not provided clearly in a study *and* you send an author query and there is no response, the final MRG should report that the study does not demonstrate that outcomes meet minimum IAA requirements.

## Examples:

(Yes) *IAA was collected at least once in each phase for each case and outcome.*

(No) *IAA was collected only in the first baseline phase for each outcome. This study does not meet minimum IAA requirements.*

(Uncertain) *IAA was collected for each case and outcome, but it is not clear if IAA was collected at least once in each phase. An author query is required to determine if this requirement was met. (This explanation would be appropriate during an initial review, before reconciliation.)*

(No) *IAA was collected for each case and outcome, but it is not clear if IAA was collected at least once in each phase. An author query was sent to determine if this requirement was met. The author did not respond, so this study does not demonstrate that it meets minimum IAA requirements. (This explanation would be appropriate to use when finalizing the MRG, after the author query time frame has passed.)*

- **Did the authors collect IAA in at least 20 percent of the sessions in each condition?**  
(row 33)

In cell B33, select “Yes” if the authors have reported IAA for at least 20 percent of the sessions in each condition. Select “No” if it is clear that this requirement is not met. If it is not clear that this requirement is met, select “Uncertain” and note in cell C33 the need for an author query.

Please draft any author questions you have on the *Author Query & Response Tab*.

## Examples:

(Yes) *The study indicates that IAA was assessed during 27 percent of the baseline sessions and 23 percent of the intervention sessions.*

(No) *The study indicates that IAA was assessed during 22 percent of the baseline sessions and 17 percent of the intervention sessions.*

(Uncertain) *The study indicates that IAA was assessed during 37 percent of the sessions. The authors did not indicate the percentage of sessions in each condition, and an author query is required to determine if this requirement was met.*

- **Does IAA meet minimum acceptable values on each relevant outcome variable?**  
(row 34)

In cell B34, select “Yes” if the authors reported minimum acceptable values on each outcome variable for each case. Select “No” if it is clear that this requirement was not met. Select “Uncertain” if it is unclear whether the study meets these requirements, and note in cell C34 the need for an author query.

In cell C34, note how the IAA is documented and provide the numerical values. If needed, draft any questions you have in the *Author Query & Response Tab*. If an author query is sent to obtain IAA information and there is no response to the IAA question(s), you should report that the study does not demonstrate that it meets minimum IAA thresholds.

*Note:* The minimum for percentage agreement is 80 percent (or .80) regardless of whether the metric is exact agreement or agreement within one. The minimum kappa or correlation is .60. IAA must meet these minimum values for each outcome across all phases or cases but not separately for each case or phase.

Example:

*(Yes) Percentage agreement was calculated by dividing the number of agreements by the sum of agreements and disagreements. IAA for cued vocalizations was 84 percent occurrence (range = 71 percent–97 percent) and 99 percent nonoccurrence (range = 98 percent–100 percent). For spontaneous vocalizations, IAA was 67 percent occurrence (range = 53 percent–67 percent) and 99 percent nonoccurrence (range = 97 percent–100 percent).*

- **Is there at least one eligible outcome that meets WWC outcome requirements?** (row 35)

Select “Yes” or “No” in cell B35. List all outcomes that meet WWC requirements with a brief description of the measure (cell C35), including the domain under which the outcome falls. Draft any questions you have on the *Author Query & Response Tab*.

Examples:

*(Yes) Both the listening comprehension and engagement outcomes meet review requirements. These outcomes fall in the reading comprehension and school engagement domains, respectively.*

*(No) The study has one eligible outcome, disruptive behavior, which falls in the problem behavior domain. The authors did not indicate if they collected IAA on at least 20 percent of all sessions and did not respond to an author query, so this study does not demonstrate that it meets minimum IAA requirements. Thus, this study does not have at least one eligible outcome that meets WWC outcome requirements. (This explanation would be appropriate to use when finalizing the MRG, after the author query time frame has passed.)*

- **Determine if the independent variable was actively manipulated:** Is the independent variable (that is, the intervention) systematically manipulated, with the researcher determining when and how the independent variable conditions change? (row 37)

In cell B37, use the drop-down menu to select “Yes” if the study provides evidence that the researcher actively manipulated the independent variable. Typically, manipulation of

the independent variable is either explicitly described in the Method section of the study or inferred from the discussion of the data collection or results.

If it is clear that the researcher did not actively manipulate the independent variable (for example, the study was conducted retrospectively using existing data), select “No.”

If any language in the study implies that the researcher did not actively manipulate the independent variable, but it is unclear, select “Uncertain” and note in cell C37 the need for an author query.

In cell C37, briefly describe how the timing and implementation of the independent variable were determined. Draft any author questions you have in the *Author Query & Response Tab*.

Examples:

*(Yes) The researcher determined in advance that the introduction of the intervention would be staggered across students. Baseline probes were collected for all students. When baseline data were stable for all students, the intervention began with the first student.*

*(No) Researchers were not involved in the decision to introduce the intervention or reintroduce the baseline condition. Teachers collected data before researchers became involved in the study, and the teachers made these decisions based on the students’ behavior.*

- **Determine if the design is free from residual treatment effects:** If there are more than two treatments in the experiment, is the design free from concerns about residual treatment effects, given the specific interventions and outcome(s) in the study? (row 39)

In cell B39, use the drop-down menu to select “Yes” if the design is free from concerns about residual effects between treatments (that is, there are not likely to be residual treatment effects).

If there are concerns about residual effects among the interventions, select “No.”

Examples:

*(Yes) There are three treatments in the experiments for all three cases in the study, but the review team leadership determined (in consultation with the content expert) that there are not likely to be residual treatment effects, given the intervention and outcomes.*

*(No) Tony’s experiment uses an ABCABC design in which intervention B always comes before intervention C for the intervention under review. In consultation with the content expert, the team determined that intervention B is likely to produce residual effects.*

- **Attempts to demonstrate an intervention effect:** Does the design include at least three attempts to demonstrate an intervention effect at three different points in time? (row 42)

Attempts to demonstrate effects occur when transitioning between conditions.

In cell B42, use the drop-down menu to select “Yes” if there are at least three attempts to demonstrate an effect at three different points in time (for at least one SCD experiment in the study).

If it is clear that this requirement was not met for any SCD experiment, select “No.”

If it is not clear whether the experiment attempts to demonstrate an effect at three different points in time, select “Uncertain” and note in cell C42 the need for an author query.

In cell C42, briefly describe the evidence for the response. Draft any questions you have on the *Author Query & Response Tab*.

Examples:

*(Yes) There are three reversal/withdrawal designs in this study. The authors used an ABA design for Louis, so there are not three opportunities to demonstrate an effect. Joy and Anton had ABAB designs, which provided at least three attempts to demonstrate an intervention effect at three different points in time.*

*(No) The authors used a multiple probe design across four students. The intervention was introduced to two students at a time—the first two students started the intervention after four baseline sessions and the second pair started after six baseline sessions. Though the study included four students with a total of eight phases, there were not three attempts to demonstrate an effect at three different points in time.*

- **Are there a sufficient number of data points in each phase?** Rows 43–46 differ based on design—the text in each row reflects the number of data points required for a design to *Meet WWC SCD Standards With or Without Reservations*. Note that review protocols may specify exceptions to the required number of data points. For example, extreme self-injurious behavior might warrant a lower threshold of only one or two data points.
- Reversal/withdrawal designs must have at least four phases per case, with at least five data points per phase to *Meet WWC SCD Standards Without Reservations* or at least three data points per phase to *Meet WWC SCD Standards With Reservations*.
- Multiple baseline and multiple probe designs must have at least six phases, with at least five data points per phase to *Meet WWC SCD Standards Without Reservations* or at least three points per phase to *Meet WWC SCD Standards With Reservations*.
- Alternating treatment designs must have at least five data points per condition and at most two consecutive data points per condition to *Meet WWC SCD Standards Without Reservations*. They must have at least four data points per condition and at most two consecutive data points per condition to *Meet WWC SCD Standards With Reservations*.

- Changing criterion designs must have at least four phases per case, with at least five data points per phase to *Meet WWC SCD Standards Without Reservations* or three or four data points per phase to *Meet WWC SCD Standards With Reservations*.

For each design, select “Yes,” “No,” or “Uncertain” to answer the questions in column A. If there is more than one SCD experiment for any type of design, provide information for each SCD experiment separately. If none of the experiments in the study use a particular design, select “Not Applicable” from the drop-down in column B (if you selected “No” for any design[s] in cells B23 through B27, the options in the corresponding drop-down lists in cells B43 through B46 will be limited to this response).

In column C, briefly describe the evidence for the response and specify the number of data points available in each phase. Draft any questions you have on the *Author Query & Response Tab*.

Examples:

(Yes) *There are at least five data points in each of the phases for Student 1 and Student 3 (this is sufficient to Meet Standards Without Reservations). For Student 2, the first baseline phase and intervention phase have four data points, but the second baseline phase has only two data points (this design does not meet standards).*

(Yes) *The study uses a multiple baseline design with six phases. Four of the phases contain at least five data points; the other two phases have only four data points. Therefore, the study can only Meet Standards With Reservations.*

(Yes) *The study uses an alternating treatment design with two interventions. Phases last at most two data points before alternating to the other intervention. There are at least five data points for each intervention (this is sufficient to Meet Standards Without Reservations).*

#### *Additional Criteria for Multiple Baseline and Multiple Probe Designs*

The following criteria apply only to multiple baseline and multiple probe designs. Multiple baseline and multiple probe studies must meet these criteria to *Meet WWC SCD Standards Without Reservations* or *Meet WWC SCD Standards With Reservations*.

- **If there are no multiple baseline and multiple probe designs in the study, select “Not Applicable” (if you selected “No” in cells B24 and B25, the options in this drop-down list will be limited to this response).**
- **Did all cases have baseline data before the intervention was administered to the first case (that is, overlapping baselines)?** There must be at least three cases with concurrent baseline phases. (row 48)

If the study meets these requirements, select “Yes” in cell B48.

If it is clear that this requirement is not met, select “No.”

If it is not clear if the requirement is met, select “Uncertain” and note in cell C48 the need for an author query.

If there are no multiple baseline and multiple probe designs in the study, select “Not Applicable” (if you selected “No” in cells B24 and B25, the options in this drop-down list will be limited to this response).

Examples:

*(Yes) All three cases overlapped—the second and third students had at least two baseline data points before the intervention was introduced for the first. The third student had five baseline data points before the intervention was introduced for the second student and three baseline data points after the intervention was introduced for the second student.*

*(No) Although the sessions in the graphs appear to be concurrent, the authors stated that they used a nonconcurrent multiple baseline design.*

- **Are the data organized in a way that allows for a vertical comparison between cases?** For plotted data, Session 1 for all cases should occur at the same time. For tabular data, the sessions should be numbered in a way that allows for easy comparison between cases of sessions that occurred at the same time. (row 49)

If the data for at least one experiment presents the timing of sessions consistently across all cases, select “Yes” in cell B49.

If it is clear that this requirement is not met, select “No.”

If it is not clear if the requirement is met, select “Uncertain” and note in cell C49 the need for an author query.

If there are no multiple baseline and multiple probe designs in the study, select “Not Applicable” (if you selected “No” in cells B24 and B25, the options in this drop-down list will be limited to this response).

Examples:

*(Yes) This is a multiple probe design, and the graphs clearly indicate when the sessions for Student 1, Student 2, and Student 3 occurred on the same day. If all students had a probe in the first three sessions, then the graphs clearly had probes when each student had a session and gaps in data when they did not.*

*(Unclear) This appears to be a multiple baseline design, but it is not clear when the sessions for each child occurred. The sessions for Gabriel, Kim, and Dorian are labeled Session 1, Session 2, Session 3, and so on, with no gaps in the data, but they appear to occur at different times (for example, Dorian’s Session 3 appears to be at the same time as Kim’s Session 6).*



- **Is training phase data present, if the design requires a training phase?**

**The design does not require a training phase if the intervention is expected to produce an effect when all of the intervention components have been implemented.**

**The design requires a training phase if the intervention is expected to produce immediate effects immediately when the intervention is implemented, even if some components are not yet in place.** (row 50)

Typically, training phases are present for an intervention that requires teaching the participants skills or practices, but the intervention will have an impact even if the training isn't complete. In some cases, researchers will choose to differentiate training data from data that takes place after the intervention is completely implemented. Training data need not be documented in cases where the researchers need to train the individuals who will be implementing the training, such as teachers, but the intervention is intended to have an effect on someone else, such as students.

Select "Yes" in cell B50 if the design collected data during a training phase and this is appropriate given the relevant research question.

Select "No" if the design did not collect data during the intervention phase and this is not appropriate given the relevant research question. The study should be rated *Does Not Meet WWC SCD Standards* because there are insufficient data to make appropriate treatment contrasts.

If the design did not collect data during the intervention phases and it is not clear whether this design is appropriate, select "Uncertain."

If the design begins data collection immediately upon implementing the intervention (whether or not it requires training) without discussing any training phases, select "Not Applicable."

If the design does not require training phase data, because the effect of the intervention is not expected until all elements of the intervention are complete, select "Not Applicable."

If there are no multiple baseline and multiple probe designs in the study, select "Not Applicable" (if you selected "No" in cells B24 and B25, the options in this drop-down list will be limited to this response).

## Examples:

(Yes) *The intervention is Self-Regulated Strategy Development, which involves multiple components. The authors do not expect to see changes in students' writing skills until after they have mastered all of the skills in the intervention.*

(No) *The authors did not collect or present data during the intervention phase, but according to the text in the study, they expected the intervention to produce differences in the children's behavior as soon as they introduced the intervention.*

- **Do cases still in the baseline phase continue baseline measurement at or after the time point when a preceding case has the first intervention data point?** (row 51)

If the study meets this requirement, select “Yes” in cell B51.

If it is clear that this requirement is not met, select “No.”

If it is not clear if the requirement is met, select “Uncertain” and note in cell C63 the need for an author query.

If there are no multiple baseline and multiple probe designs in the study, select “Not Applicable” (if you selected “No” in cells B24 and B25, the options in this drop-down list will be limited to this response).

#### *Additional Criteria for Multiple Probe Designs*

The following criteria apply only to multiple probe designs. Multiple probe design studies must meet these criteria to *Meet WWC SCD Standards Without Reservations* or *Meet WWC SCD Standards With Reservations*.

- **Do initial baseline sessions for each case overlap vertically? Within the first three baseline sessions, the design must include three consecutive probe points for each case to *Meet WWC SCD Standards Without Reservations*. The design must include at least one probe point for each case to *Meet WWC SCD Standards With Reservations*. The probe points need not be in exactly the same session in all cases, so long as there is at least one probe point in the first three sessions for each case.** (row 53)

If the study meets either of these conditions, select “Yes” in cell B53.

If it is clear that this requirement is not met, select “No.”

If it is not clear whether either of these conditions is met, select “Uncertain” and note the need for an author query.

If there are no multiple probe designs in the study, select “Not Applicable” (if you selected “No” in cell B25, the options in this drop-down list will be limited to this response).

In cell C53, briefly describe the evidence for the response. Draft any questions you have on the *Author Query & Response Tab*.

Examples:

(Yes) *The baseline sessions overlap and there are three probe points in the first three sessions for each case.*

(Yes) *There are three probe points in the first three sessions for the first and third cases. There is only one probe point in the first three sessions for the second case. Therefore, the highest rating this study can receive is Meets WWC SCD Standards With Reservations.*

(No) *The first probe for the third case does not occur until the fifth session. The study cannot meet WWC SCD Standards as a multiple probe design.*

- **Are probe points available just before introducing the independent variable? Within the three sessions just before introducing the independent variable, the design must include three consecutive probe points for each case to *Meet WWC SCD Standards Without Reservations* or at least one probe point for each case to *Meet WWC SCD Standards With Reservations*.** (row 54)

If the study meets either of these conditions, select “Yes” in cell B54.

If it is clear that this requirement was not met, select “No.”

If it is not clear whether either of these conditions was met, select “Uncertain” and note in cell C63 the need for an author query.

If there are no multiple probe designs in the study, select “Not Applicable” (if you selected “No” in cell B25, the options in this drop-down list will be limited to this response).

In cell C54, briefly describe the evidence for the response. Draft any questions you have on the *Author Query & Response Tab*.

Example:

(Yes) *The data include three probe points just before the introduction of the intervention in two of three cases. The third case has only one probe point just before the introduction of the intervention. The highest rating this study can receive is Meets WWC SCD Standards With Reservations.*

- **Does each case not receiving the intervention have a probe point in a session in which another case either first receives the intervention or reaches the prespecified intervention criterion?** (row 55)

If the study meets either of these conditions, select “Yes” in cell B55.

If it is clear that neither requirement was met, select “No.”

If it is not clear whether either of these conditions is met, select “Uncertain” and note in cell C55 the need for an author query.

If there are no multiple probe designs in the study, select “Not Applicable” (if you selected “No” in cell B25, the options in this drop-down list will be limited to this response).

In cell C55, briefly describe the evidence for the response. Please draft any questions you have on the *Author Query & Response Tab*.

Examples:

*(Yes) The second case in the series has a probe point in the session in which the first case receives the intervention. The third case has probe points in the sessions in which the first and second cases receive the intervention.*

*(No) The third case in the series does not have a probe point in the session in which either the first case receives the intervention or reaches the intervention criterion. This study does not meet WWC SCD standards.*

### Other Study Design Issues

The following criteria should be completed for all SCD studies, regardless of study design.

- **Is the study free of other data or analytic issues that would affect the rating?** (row 57)

In cell B57, select “Yes” if there are no other issues about the analysis or data that an earlier entry did not capture. If there are issues, select “No.” Summarize any issues in cell C57.

- **Is there any attrition or sample loss not otherwise noted in the SRG?** (row 58)

Select a response using the drop-down menu in cell B58.

If the case is a classroom or other group of students, attrition could affect the study results. Attrition also can occur in studies using individuals as cases if the individual fails to attend all of the planned sessions. Please use cell C58 to describe any attrition not captured in the data points discussion.

- **What is the highest rating of an experiment in the study, given current information?** (row 59)

Select a rating using the drop-down menu in cell B59.

If you select *Meets WWC SCD Standards With (or Without) Reservations*, in cell C59 select “The study *Meets WWC SCD Standards With or Without Reservations*. An explanation for the highest rating is below.” In cell C60, briefly describe the contrast(s)

and rating(s) that meet those standards. You will need to continue the study review in Stage 3.

If the study is rated *Does Not Meet WWC SCD Standards* based on the information provided, select the appropriate disposition code in cell C59 and provide an explanation for the disposition code in cell C60. Copy and paste cell C60 as many times as needed to capture the exact disposition code (selected drop-down from menu in cell C59) for each comparison that could be rated *Does Not Meet WWC SCD Standards*.

The *Does Not Meet WWC SCD Standards* disposition codes are as follows:

... no graphical illustration or table of data is provided. This covers scenarios in which the SCD data are not presented in the study graphically or in tabular format.

... there are insufficient data to demonstrate an intervention effect. This covers scenarios in which (a) there are an insufficient number of attempts to demonstrate the intervention effect, (b) there are not enough data points to demonstrate an intervention effect, or (c) a multiple baseline or multiple probe design fails to meet the additional criteria for these types of designs.

... the eligible outcomes do not meet WWC requirements. This covers scenarios in which (a) outcomes were overaligned with the intervention, (b) outcomes were determined not to be sufficiently valid or reliable, or (c) IAA data were not reported or did not meet minimum thresholds.

... the measures of effectiveness cannot be attributed solely to the intervention. This covers scenarios in which (a) there was a confounding factor (for example, interventionists were confounded with particular phases or conditions), (b) there was reason to believe that there were residual treatment effects present in the study, or (c) the effects of the intervention of interest were reported only in combination with other interventions.

... conditions for manipulation of the independent variable are not determined by the researcher. This covers scenarios in which the researcher was not involved in deciding the timing of changes between phases.

If an author query is needed to determine the final rating, indicate what response would be needed. The author query should be referenced in the appropriate row above and the question drafted on the *Author Query & Response Tab*.

In row 60, provide a full explanation for the rating and any disposition code.

**Examples:**

*The study is rated Meets WWC SCD Standards Without Reservations. The study used one multiple baseline design experiment with three participants and six phases. The data points per phase ranged from six to nine. The outcomes meet review requirements.*

*The study is rated Meets WWC SCD Standards With Reservations. It included one reversal-withdrawal design experiment (ABAB). The first baseline phase had only four data points; the remaining three phases each had six data points. The outcome meets review requirements.*

*The study is rated Does Not Meet WWC SCD Standards due to insufficient data. It was a reversal experiment that used an ABA design. It did not have the minimum required number of phases to meet standards.*

*The study is rated Does Not Meet WWC SCD Standards due to insufficient data. It was a multiple probe experiment with three participants. Though the study had sufficient data points per phase to meet standards, Participants 2 and 3 lacked a probe point just before the intervention was introduced to them.*

*The study is rated Does Not Meet WWC SCD Standards because the eligible outcomes do not meet WWC requirements. There were three separate ABAB experiments in the study (for three separate students), and none of the experiments demonstrated minimum IAA thresholds, following an author query. Two of the designs also had insufficient data because the baseline phases had only two data points.*

*The study is rated Does Not Meet WWC SCD Standards due to insufficient data. There were four reversal-withdrawal experiments in the study for four separate students. Two of the experiments were AB designs (and thus did not provide at least three opportunities to demonstrate an effect). The other two designs collected IAA data only during the intervention phases (and thus did not meet minimum IAA requirements). The intervention report will use the insufficient data disposition code.*

- **What is the rating and disposition code for each SCD experiment, given the current information?** (rows 61 and 62)

Copy and insert rows 61 and 62 for each SCD experiment in the study so that each SCD experiment has at least one rating or disposition code and explanation. For example, if a study has one reversal-withdrawal experiment for Marcela that does not meet standards due to “insufficient data” (because it was an ABA design), and a second reversal-withdrawal experiment for Alejandra that does not meet standards because “no graphical illustration of evidence is provided,” you should copy and insert rows 61 and 62 and provide separate ratings, disposition codes, and explanations for Marcela and Alejandra (that is, there are two rows for each). In cell C62, enter information specific to the experiment, including the participant’s name (or identification number), design, and

outcome, so that the disposition code(s) and explanation clearly refer to a specific experiment.

If a single SCD experiment has multiple disposition codes, copy and insert row 61 so that you can select each relevant disposition. For example, if there is a reversal-withdrawal experiment for Sam that does not meet standards because the “eligible outcome does not meet WWC requirements” (due to not demonstrating minimum IAA thresholds) and because it has insufficient data (it is an ABA design), you should copy and insert row 61 for Sam’s experiment so that you can select both reasons this experiment does not meet standards.

In cell C62, enter the explanation for the disposition code(s) selected in row 61 (and the other copied and inserted rows). Do this separately for each experiment.

If you insert multiple dispositions for one SCD experiment (in row 61), make sure to provide explanations for each of the disposition codes in row 62.

If there is only one experiment in the study, you should still complete rows 61 and 62. You can use text from the overall study rating but add experiment-specific information in cell C62. You should also include a note in cell C62 to indicate that there is only one experiment in the study.

Examples:

(cell C62) *Rosa (School Engagement). Rosa’s ABA reversal-withdrawal experiment does not meet standards because there are only two opportunities to demonstrate an effect.*

(cell C62) *Victor (School Engagement). Victor’s reversal-withdrawal experiment does not meet standards because there is no graphical presentation of the data.*

(cell C62) *Sam (Problem Behavior). Sam’s ABA reversal-withdrawal experiment does not meet standards for two reasons. First, there are only two opportunities to demonstrate an effect. Second, the problem behavior outcome does not meet WWC requirements because it does not meet the minimum IAA threshold for percent agreement. (Note, for this experiment, you would copy and insert row 61 so there is space for two disposition codes: [a] ... there are insufficient data to evaluate the attempts to demonstrate an intervention effect and [b] ... the eligible outcomes do not meet WWC requirements.)*

- **Document any information needed to complete this review.** (row 63)

Summarize additional information needed to complete the review, including how the overall study rating and/or ratings for individual SCDs could change based on the results of an author query. Summarize all missing information here, even if the study or design does not meet standards for another reason.

If an author query is needed to determine the final rating, indicate what response is needed. Refer to the author query in the appropriate row above and the question(s) drafted on the *Author Query & Response Tab*.

For example, if a study does not clearly show that IAA was collected in each phase, but all other criteria for meeting standards are in place, indicate that an author query will be sent to clarify whether IAA was collected in each phase, but that given the information currently available, the study is rated as *Does Not Meet WWC SCD Standards*. However, pending an author query, the study might be rated *Meets WWC SCD Standards (With or Without) Reservations*.

If a study does not clearly show that IAA was collected in each phase, and the study uses an AB design that does not meet standards due to insufficient attempts, indicate that an author query will not be sent to clarify whether IAA was collected in each phase because the study is already rated as *Does Not Meet WWC SCD Standards*. In this case, do not use the disposition code “eligible outcome does not meet WWC requirements” (due to not demonstrating minimum IAA thresholds), because you did not give the authors a chance to provide this information in an author query; instead, you would only use the disposition code “insufficient data” (because it is an AB design).

When finalizing the MRG (after the author query period has passed), clearly document what information you received or did not receive in this cell.

### ***Stage 3: Study Details***

Stage 3 of the SRG summarizes the key findings and a broad description of the study design and intervention. Fill out the beginning part of this section through the contextual variables, regardless of the overall study disposition. Ideally, this section should be written so that the text for each subsection can be directly pasted into the appropriate appendices for an intervention report or individual study review. Therefore, do not use any text that was directly copied from the study in the descriptions of the study details in Stage 3. Please fill out this section with regard to the whole study, not only those experiments that met standards.

- **Did the authors present effect sizes? If so, how were they computed?** (row 66)

If the study authors present effect size estimates, select “Yes” in cell B65 and indicate how those effect sizes were computed in cell C66 of the *Main Tab* (regardless of whether impact estimates are provided in other metrics). If the authors do not report effect sizes, select “No.” In particular, note whether the author-reported effect size differs from the WWC-calculated effect size and provide information that explains why (for example, the authors computed a nonoverlap effect size or a within-case standardized mean difference).



Examples:

1. (Yes) *The authors do not report effect sizes.*
2. (Yes) *The authors report effect sizes; however, they reported the nonoverlap of all pairs (NAP), a nonoverlap effect size estimate.*

- **Are any findings reported for eligible subgroups as defined in the review protocol?** (row 67)

In cell B67, select “Yes” if the study provides impact estimates for any of the subgroups outlined in the protocol under *Eligible Populations*; select “No” otherwise.

In cell C67, briefly describe which groups are analyzed and whether those analyses *Meet WWC SCD Standards*. (See the instructions on completing the *Review Tab* for more information.)

- **Number of students** (row 69)

Enter the number of students within the analytic sample in cell B69. If the number of students in the analytic sample varies by subgroup or outcome, provide details in cell C69.

- **Delivery method** (row 70)

Select one of the options from the drop-down menu in cell B70: *Individual, Small group, Whole class, School, or Mixed delivery method*. If more than one delivery method was used, list all methods in cell C70.

- **Program type** (row 71)

Select one of the options from the drop-down menu in cell B71: *Curriculum, Practice, Supplement, School-level, Teacher-level, Policy, or Mixed program type*. If more than one type was used, list all types in cell C71.

- **Type of school** (row 72)

Select one of the options from the drop-down menu in cell B72: *Public, Private, Parochial, Charter, or Mixed type of school*. If more than one type of school was included, list all types in cell C72.

- **Type of classroom** (row 73)

Select *General, Inclusion, or Both general and inclusion* from the drop-down menu in cell B73.

- **Grades or ages** (row 74)

Select one of the options from the drop-down menu in cell B74: *Pre-K, Kindergarten, 1–12, Postsecondary, or Mixed grades/ages*. If the *Mixed grades/ages* option is selected, list additional details about the grades/ages covered in cell C74.

- **Geographical setting** (row 75)

Select one of the options from the drop-down menu in cell B75: *Rural, Suburban, Urban, or Mixed geographical setting*. If more than one geographical setting was included in the study, list all settings in cell C75.

- **Location** (row 76)

Select one of the options from the drop-down menu in cell B76: *West region, Northeast region, Midwest region, South region*, and an individual listing of all 50 U.S. states, as well as *Mixed locations*. If the study took place in more than one region/location not covered by the specific listings and you selected *Mixed locations*, list all locations in cell C76.

- **Percent Free or Reduced-Price Lunch (FRPL); English Language Learner; IEP (Individualized Education Program)/special education; White; Black or African American; American Indian or Alaskan Native; Asian; Native Hawaiian or Other Pacific Islander; Other; Hispanic; Non-Hispanic; Male; Female** (rows 77–89)

Enter a percentage in column B for each contextual variable. Record “nr” for any contextual variables not reported in the study. Briefly describe any important context in column C.

*If the study passes Stage 2... (row 90)*

Complete the remaining sections of the SRG only for those studies which are rated as *Meets WWC SCD Standards With or Without Reservations*.

Complete *Data Tab* if applicable. (row 91)

Click on cell A91 to be taken to the *Data Tab*. Fill out the *Data Tab* as completely as possible (instructions begin on page 36).

Complete Effect Size Section of *Review Tab* (columns N–R) if applicable. (row 92)

Click on cell A92 to be taken to the WWC-computed findings sections of the *Review Tab* (instructions begin on page 33).

- **Setting of the study (for example, location, classrooms, courses, schools)** (row 94)

Describe the location of each SCD experiment. Include the names of states or U.S. regions when possible. Provide any details regarding the specific location of the sessions, such as the regular classroom or special locations for the interventions.

Example:

*The study was conducted in an elementary school in a suburban area in the northwestern United States. The intervention for Ben took place in his regular classroom. The intervention for Gilberto took place in the hallway, right outside of his regular classroom.*

- **Study design** (row 95)

Summarize the study design, including how the introduction of the independent variable was determined.

Example:

*This study used a reversal design to identify the effect of an intervention designed to increase on-task behavior. The ABAB design enabled the researcher to observe the student under baseline conditions before introducing the intervention. The intervention was introduced after five sessions of stable baseline performance were observed. It was then withdrawn after eight sessions. It was reintroduced after an additional eight sessions in the second baseline phase.*

- **Sample sizes (for example, students, classrooms, teachers, schools)** (row 96)

Summarize the participants who are from an eligible population and are from designs that met standards (with or without reservations). Report if there are any cases or students that are ineligible for review or do not meet WWC SCD standards.

Example:

*Four eligible students participated in the study. One additional student participated in the study but was ineligible for review.*

- **Sample characteristics in protocol (for example, race, gender, FRPL) (row 97)**

Describe the characteristics of participants who are eligible and have designs that meet standards (with or without reservations). Include any pseudonyms used in the publication. Include racial and ethnic information for the sample, if available.

Example:

*James was a 10-year-old white male diagnosed with a learning disability in reading as well as attention-deficit/hyperactivity disorder (ADHD). Gustavo was an 11-year-old Latino male whose primary language was Spanish. He was diagnosed with learning disabilities in math, reading, and language arts. Esteban was an 11-year-old Latino male whose primary language was Spanish. He was diagnosed with learning disabilities in math, reading, and language arts, and with ADHD. Claire was an 11-year-old African American female diagnosed with learning disabilities in math, reading, and language arts.*

- **Intervention condition as implemented in the study (including number of days/weeks/months, number of sessions, time per session) (row 98)**

Summarize the intervention(s) in sufficient detail to help readers understand what makes this intervention similar to or different from others. The level of detail should be similar to what an introductory section of a typical impact evaluation report would provide. It should include the length and the dosage of the intervention, as well as information about the content and delivery of the intervention. Note any intervention differences between cases or settings. This description should be about the intervention as used in this study, not as described by the developer or in ideal conditions.

Example:

*During the intervention phases, students used a checklist to prompt each step of the meal preparation process. Each session took between 30 and 45 minutes, depending on how successful the students were at completing the checklist. The instructor started each session by showing the student a picture of the meal to be prepared and instructed the student to use the checklist to aid in the preparation process. The student then completed each step, checking steps off as they were completed. If the student failed to complete a step correctly, the instructor prompted the student to refer to the checklist. If the student still failed to complete the step correctly, the instructor completed the step outside of the student's view. The instructor then prompted the student to proceed to the next step. Intervention sessions continued until the student completed all steps correctly without having to repeat any of them. Students took between five and nine sessions to complete all steps correctly.*

- **Baseline condition as implemented in the study** (row 99)

Indicate procedures used in the baseline (or comparison) condition. Clarify whether the comparison condition was a particular alternative intervention and, if so, name the intervention and provide a brief description if the study provides that information.

Example:

*The study used a reversal-withdrawal design. Observation sessions took place three times per week, each session lasting 30 minutes. Teachers were observed for two weeks, for a total of six sessions. During the baseline/withdrawal condition, there was no formal behavior management system in place in the classroom. Observations revealed that the teacher occasionally praised students or called out to the class that they were doing well. The consequences for inappropriate behavior were inconsistent.*

- **Any maintenance phases** (row 100)

Describe any maintenance phases (after the intervention has ended), including the timing and conditions.

Example:

*For Aaron, maintenance data were collected seven weeks after intervention completion. At that time, Aaron's teacher had discontinued the use of the checklist and did not give him opportunities to choose peers for group work.*

- **Describe all eligible outcomes reported and how they were measured.** (row 101)

Describe all of the outcomes within relevant domains examined in the study and identify which of those outcomes are eligible based on the criteria specified in the protocol under *Types of Outcomes to Be Included*. In addition, indicate how each eligible outcome was measured (if it is not self-explanatory from the name of the outcome), how it was collected (if relevant), and how it can be interpreted (the scale of the measure). Use quotes around any text that comes directly from the original study.

Example:

*There are two eligible outcomes:*

***Inappropriate behavior*** (problem behavior domain) was defined as the frequency of inappropriate behaviors during a 10- to 15-minute session, using 10-second partial interval recording. Todd's inappropriate behavior was defined as "screaming, throwing, hitting, crying, yelling, and refusing to move." Guillermo's inappropriate behavior was defined as "forcefully hitting others." Eli's inappropriate behavior was defined as "leaving the classroom, turning over furniture, and physical aggression toward staff."

***Time on task*** (school engagement domain) was defined as the percentage of time the student was on task and engaged in seatwork, during a 10- to 15-minute session. This outcome was defined similarly for all three students.

- **For any outcome that met standards with or without reservations, was visual analysis described in the text? If yes, provide details regarding the visual analysis and any conclusions about the functional relation between the intervention and the outcome(s).** (row 102)

Select "Yes" or "No" in cell B102. If "Yes," provide details in cell C102. In the case of this description, it is appropriate to paste direct quotes of material from the text into the cell. Ensure that all direct quotes are enclosed in quotation marks and are clearly indicated as direct quotes. In the cases where the visual analysis includes both experiments or phases that met standards and experiments or phases that did not, include the total visual analysis and clarify which cases met standards and which did not. This could be the case in a visual analysis of a reversal/withdrawal design where only a subset of phases met standards, or a multiple baseline/multiple probe where only a subset of cases met standards.

For the important elements of visual analysis, please review the WWC SCD technical documentation, pages 17 through 21 ([found here](#)).

- **Are there outcomes that do not meet review requirements? If yes, provide the domain and a brief description of the reason.** (row 103)

Select “Yes” or “No” in cell B103. In cell C103, list all eligible outcomes and why they do not meet review requirements. If outcomes vary by case or SCD experiment, specify which outcomes go with each case or SCD experiment.

Example:

*(Yes) The study did not present data for the “on-task behavior” outcome, so this outcome does not meet WWC standards.*

- **Are there any outcomes that are not eligible for review? If yes, provide a brief description and the reason.** (row 104)

Select “Yes” or “No” in cell B104. In cell C104, list the outcomes that are not eligible for review, along with a reason for ineligibility. If outcomes vary by case or SCD experiment, specify which outcomes go with each case or SCD experiment.

Example:

*(Yes) The social cooperation outcome is not eligible for review under the Students With Specific Learning Disabilities protocol. It does not fall under any of the topic area’s academic domains.*

- **Support for implementation** (row 105)

Indicate both the staff training and technical assistance conducted to support the implementation of the intervention (as evaluated in the study).

Example:

*The students’ regular classroom teacher conducted the instruction in both the baseline and intervention conditions. The teacher met privately with the researcher before the start of the study and received brief training on how to use the instructional videos as part of the intervention.*

- **Implementation fidelity** (row 106)

Describe if and how implementation fidelity was assessed, and any ratings provided by the authors.

Example:

*The instructor followed a detailed script when implementing both the baseline and intervention procedures. Implementation fidelity was measured simultaneously with IAA. Independent observers determined that baseline and intervention procedures were implemented correctly in 100 percent of sessions for both participants.*

- **Discuss other SCD experiments presented in the study that were not part of the review.** (row 107)

Describe any additional SCD experiments presented but not already covered or described in the current SRG. Explain why the review did not include these SCD experiments.

Example:

*The authors also presented an SCD experiment used to measure the effect of a different intervention, called Patch. This SCD experiment is not relevant to the current review effort because it focuses on functional behavioral assessment and not Patch.*



### ***Review Tab of the SRG***

Each row corresponds to a research question addressed by an SCD. The reviewer should represent each SCD experiment (with an eligible outcome) on a separate row. If a study contains data on more than one outcome for a participant, each outcome should receive a separate row.

For example, use a separate row to represent each reversal-withdrawal design presented in a study. If there are three separate reversal-withdrawal designs for three participants, each participant would get their own row. If there are separate reversal-withdrawal designs for three different outcomes (across one participant), each outcome would get its own row.

Represent multiple baseline and multiple probe designs on one row, such that the row represents the entire SCD experiment. Again, if the multiple baseline or multiple probe design offers data on more than one outcome, each outcome should get a separate row.

For alternating treatment designs, each comparison (that is, deemed relevant to the review) should have its own row.

Finally, columns W, X, and Y are columns for the finalized review guide produced by the reconciler or the leaders of the review. These columns are used to record information about the reliability of the data extracted from SCD plots, if tabular data are not available in the study or from an author query.

### **Measures and Domain Characteristics Section of the Review Tab**

General instructions:

- Include only outcomes that fall within one of the eligible domains for the review.
- If you are uncertain whether a particular outcome measure falls under one of these domains, check for additional guidance that your team may provide on outcome measures.
- If results are presented separately for different experiments, you will need to have a separate row for each. If reliability information differs across the samples, make sure it is accurately reflected in the table.
- **Measure** (column A)  
Indicate the name of the outcome or test, exactly as it was specified in the article or report.
- **Domain** (column B)  
Indicate the eligible domain(s) for the outcome.
- **Design** (column C)

Select the design corresponding to the measure you are reporting on. For multiple baseline and multiple probe designs, you will need to specify whether it is a design across several units (“multiple baseline across participants” or “multiple probe across participants”) or a design within a single unit (“multiple baseline across settings or outcomes” or “multiple probe across settings or outcomes”).

- **Sample** (column D)

List the participants (sample size) for the experiment in this column (for example, enter “Christina [1]” for a reversal/withdrawal experiment focused on one participant; “Low-performer group [10]” for a reversal/withdrawal experiment focused on a group of 10 low-performing students; or “Julio, Susana, and Adam [3]” for a multiple baseline experiment across three participants).

- **Outcome period** (column E)

Indicate any relevant information about the timing of intervention or measures (for example, baseline and intervention).

- **Binary** (column F)

Select “Yes” from the drop-down menu if the outcome is a binary variable for which the underlying construct is a yes/no answer, such as “ever graduated” or “retained in grade.” A “Yes” response indicates that there is not an underlying distribution of the variable. Select “No” otherwise.

- **Standardized test** (column G)

Select “Yes” from the drop-down menu if the test is a standardized test. The score should be from the full test or an established subscale to be considered a standardized test.

- **Face validity** (column H)

Select “Yes” from the drop-down menu if the measure appears to be a reasonable measure; select “No” if you see an obvious problem with the measure, and indicate your concerns in column P.

- **Not overaligned?** (column I)

Select “Yes” from the drop-down menu if you have no concerns that the measure may be overaligned with the intervention.

Select “No” from the drop-down menu if you have concerns that the measure may be overaligned with the intervention.

Measures that are closely aligned or tailored to the intervention are likely to demonstrate larger effect sizes than those that are less closely aligned with the intervention. An example of overalignment is if the measure includes some of the same materials (such as

specific reading passages) that are used in the intervention or administered to the intervention group as part of the intervention. Explain any concerns in column P.

- **Consistent measure and data collection process?** (column J)

Select “Yes” from the drop-down menu if the intervention and baseline conditions collected the same measure in a similar manner; select “No” if it is clear that the intervention and baseline conditions collected outcome data in a different manner, potentially in a way that could lead to differences in average outcomes by condition.

Explain any concerns in column P.

- **Meets IAA requirements (see *Main Tab*)?** (column K)

Select “Yes” from the drop-down menu if the IAA requirements are met (*Main Tab* rows 32–34). If the experiment does not demonstrate that the requirements are met, select “No” from the drop-down menu.

Explain any concerns in column M.

- **Meets WWC requirements?** (column L)

Based on the answers in the previous columns and the standards established in the review protocol, select “Yes” from the drop-down menu if the outcome meets all of the requirements; select “No” otherwise.

If you select “No,” ensure column N includes details on why the measure does not meet WWC requirements.

- **Notes or concerns about the measure** (column M)

Summarize any concerns you have about this measure. In particular, if the measure does not meet the requirements for inclusion in this review as noted in column L, indicate why.

If a citation is provided for a standardized test or subscale, it may be helpful to note that here.

If the tabular data were available, either via the primary study or from an author query, note this in column U.

Now, go back and enter a decision in cell C35 on the *Main Tab*. (Is there at least one relevant outcome that meets review requirements?)

### **Effect Size (WWC-Calculated Findings) Section of the Review Tab**

For each experiment that is rated *Meets WWC SCD Standards With or Without Reservations*, and for which there is a design-comparable effect size, the reviewer should record the design-comparable effect size (estimated using software specified by the leaders of the review), the

standard error of the design-comparable effect size estimate, the degrees of freedom for the effect size, the  $p$  value for the effect size, the autocorrelation estimate, and any nondefault modeling choices that were made when estimating the effect size.

- **Design-comparable effect size estimate** (column N)

Record the design-comparable effect size estimate (if available) in this column.

- **Change sign of effect for meta-analysis?** (column O)

In a systematic review, the WWC prefers for desirable changes to be represented by positive effect sizes. In some cases, the calculated effect size might have a negative effect for a desirable outcome, or a positive effect for an undesirable outcome. For instance, in a study intended to reduce problem behaviors, a *reduction* in problem behaviors is considered desirable, but would correspond to a *negative* effect size estimate in terms of the observed outcome. In a review of an intervention to reduce problem behaviors, the WWC would change the sign of the effect size to be positive to reflect a desirable change in the behavior.

If the sign of a negative effect size estimate should be changed to positive so that it reflects a desirable change, or the sign of a positive effect size should be changed to negative to reflect an undesirable change, select “Yes” in this column in the row corresponding to the effect size whose sign should be changed. If the direction of the effect size should remain unchanged, select “No” in this column in the row corresponding to the effect size.

- **Standard error** (column P)

Record the standard error of the design-comparable effect size estimate in this column.

- **Degrees of freedom** (column Q)

Record the degrees of freedom for the design-comparable effect size in this column.

- **WWC-estimated  $p$  value**

This column will auto-populate a  $p$  value based a two-tailed t-test for the design-comparable effect size when columns N, P, and Q are filled out for a given row.

- **Autocorrelation estimate** (column S)

Record the autocorrelation estimate (if available) to the design-comparable effect size in this column.

- **Note any deviations from default modeling guidance** (column T)

If you have deviated from the default guidance for design-comparable effect size estimation specified in the *WWC Procedures Handbook*, either because of guidance from the review protocol or the leaders of the review, please record those deviations in this column.

**Inter-reviewer reliability section only needs to be filled out during reconciliation.**

The final three columns (columns U-W) are available to record any necessary details for reconciling data extracted from plots in the primary study when the data are not available in tabular format.

- **Reliability method** (column U)

Record the type of reliability estimate (such as the ICC) used to estimate the reliability of extracted outcome data for the two reviewers.

- **Reliability estimate** (column V)

Record the reliability estimate for the extracted outcome data.

- **Notes about reliability/extracted data** (column W)

Note here any other issues or concerns about the reliability or the data extraction process. Note if the data were available in tabular format (either from the primary study or through author query) rather than extracted.

### ***Data Tab of the SRG***

Data from experiments that *Meet WWC SCD Standards With or Without Reservations* should be captured in the *Data Tab*. The *Data Tab* is organized with individual rows corresponding to a single observation. Five columns come prespecified, but more columns can be added as needed.

- **Participant** (column A)

Specify the participant or unit's name, pseudonym, or identifier exactly as it is provided in the original study.

- **Behavior, Outcome, or Context/Setting** (column B)

Record the name of the behavior, outcome, or context/setting. This should be the name of the outcome of interest for the experiment. In the case of reversal/withdrawal, changing criterion, and alternating treatment designs with multiple experiments, this is all likely to be the same or a similar behavior. For multiple baseline designs across contexts or behaviors, this will typically be the same behavior in different contexts.

- **Phase Name** (column C)

Record the phase the data point was drawn from, as it was described in the data or plot. This might be a phase letter such as "A," "B," or "C"; it might be the name of the intervention in an alternating treatment design; or it might be generic "baseline" and "treatment" descriptors.

- **Session Number (x-axis)** (column D)

Record the session ordering. Session numbers should generally be whole numbers that maintain any vertical comparisons present in the original data or plots. For instance, in a multiple baseline design where all cases had data gathered on the same day, the first session for each case would be numbered as "1."

- **Outcome Value (y-axis)** (column E)

Record the actual value for the outcome for this data point. If the data are displayed graphically, this will be the value for the data point on the *x*-axis.

*Note:* It should be possible to identify individual cases by the unique combination of the **Participant** and **Behavior, Outcome, or Context/Setting** columns. If identifying individual cases requires recording an additional variable, start another column with a heading in row 1 and record values for each data point in that column.

The reconciler or the leader of the review that finalizes the review guide for the study should ensure that copies of the data tab for both reviewers are included in the final SRG if the data had to be extracted directly from the study plots.

### ***Plots Tab of the SRG***

This tab will only be used if the data are *only* available in graphical format—that is, the data are not available in tabular format and an author query for data did not receive a response. In that case, reviewers will need to extract the data from the plots. When data are extracted from plots, please also copy and paste the plots used to extract the data for the *Data Tab* of the SRG. Ensure that the plots are labeled in such a way that they are easily matched to their corresponding lines in the *Review Tab* and *Data Tab*. This can take the form of labels included in the images or text labels in the spreadsheet.

The reconciler or leader of the review that finalizes the review guide for the study should ensure that copies of the *Plots Tab* for both reviewers are included in the final SRG if the data had to be extracted directly from the study plots.

### ***Author Query and Response Tab of the SRG***

Reviewers should draft potential author query questions as part of their review of a study. The MRG should capture the final set of questions as sent to the author. It also should document whether (and when) a response was received, and the response text.

- **Date author query sent** (cell B2)

Enter the date the author query was sent.

*Note:* Check with review leadership to determine who should send the author query.

- **Response received? (Yes/No)** (cell B3)

Enter “Yes” or “No” to document whether any response was received.

- **Date of response** (cell B4)

Enter the date the team received the response(s).

- **Mode of response (for anything other than an e-mail)** (cell B5)

Enter the particular mode(s) by which the team received the response(s), in particular documenting if the response was received by some mode other than e-mail.

- **Author query questions and answers**

Save questions and responses in rows 6–11. Copy and insert additional rows as needed. If the author query included a table, enter the values the author provided in any table shell (or copy and paste populated tables, if appropriate).

### ***Summary Tab of the SRG***

This worksheet captures high-level information about the study, rating, outcomes, and findings. Many of the cells in this worksheet will populate automatically based on the data entered in the *Main Tab* and *Data Tab* of the SRG. These cells are shaded gray. They include summary-level and experiment-specific information:

#### **Study-specific entries**

- Study ID (cell G2)
- Full Citation (cell H2)
- Review Date (cell I2)
- Standards Version (cell J2)
- Review Protocol and Version (cell K2)

#### **Experiment-specific entries**

- Measure (column A)
- Domain (column B)
- Design (column C)
- Sample (column D)

In addition to the prefilled cells, reviewers will have to add data on the age(s) or grade(s) associated with each case, the WWC design rating for each experiment, and author-reported findings. These cells are shaded green.

- Enter the age(s) or grade(s) associated with the experiment's sample in column E.
- Use the drop-down menu to select the WWC rating associated with each SCD experiment in column F.