Systematic Reviews of Single-Case Research: Aims, Challenges, and Good Practices

James E. Pustejovsky
pustejovsky@wisc.edu

Daniel D. Drevon
drevo1dd@cmich.edu

September 26, 2025

Outline

1. Systematic review and synthesis of SCDs
2. Aims & scope of systematic reviews
3. Challenges of conducting intervention-focused reviews
4. Good practices and recommendations

Systematic Review, Research Synthesis, & Meta-Analysis

- **Systematic review:** A literature review that uses systematic search techniques and inclusion criteria to identify a body of relevant literature. Systematic review methods aspire for
  - Transparency
  - Reproducibility & Replicability
  - Comprehensiveness
  - Impartiality (mitigation of personal biases / preconceptions)

- **Research synthesis**: the systematic integration of empirical research for purposes of drawing generalizations (Cooper & Hedges, 2009).

- **Meta-analysis**: statistical methods that support research synthesis, especially methods for combining results from a collection of studies.

Disciplines that rely on research synthesis

- Medicine (Cochrane Collaboration)
- Education (What Works Clearinghouse)
- Psychology
- Social policy (justice, welfare, public health, etc.)
- Economics, international development
- Physical sciences

AIMS & SCOPE

Many possible aims of a systematic review
A non-exhaustive list:

What research is available in a field / on a topic?

How has past research been conducted on a topic?

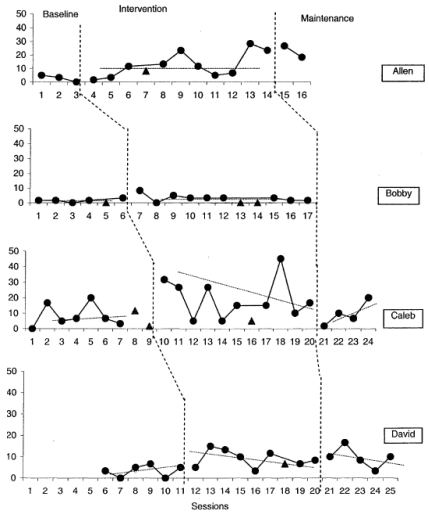
What is the prevalence of {some phenomenon} in {some population}?

What are the effects of {an intervention / practice} on {some outcome} for {some population}?

What type of review is needed?

What research is available in a field / on a topic?	Scoping review / evidence map
How has past research on a topic been conducted?	Methodological feature / quality review
What is the prevalence of {some phenomenon} in {some population}?	Meta-analysis of prevalence
What are the effects of {an intervention / practice} on {some outcome} for {some population}?	Synthesis / meta-analysis of intervention research

Synthesis of intervention research for informing evidenced-based practice



- Some SCDs are designed to provide evidence about intervention effects ***for individual participants***.
- But single SCDs provide limited basis for ***generalization*** to other participants or contexts.
- Combining evidence from multiple studies can provide a firmer basis for generalization about effects of intervention.
  - ***Summarize*** findings across studies
  - Describe extent of ***variation (heterogeneity)*** in effects
  - Identify ***predictors*** and ***boundary conditions*** of effectiveness

Defining the Scope of a Review

- Scope of review is defined through operational inclusion criteria

- A helpful acronym: PICOS+D
  - Participants (possibly only subset of participants from a study)
  - Interventions (possibly only certain intervention phases)
  - Comparisons (possibly only certain comparison phases)
  - Outcomes (possibly only some DVs)
  - Settings (generalization? maintenance?)
  - Designs / methods (design standards, quality criteria)

- See Pustejovsky and Ferron (2017), Ledford and colleagues (2022), Becker (2017), Moeyaert (2019) for more detailed discussion.

Ledford & Pustejovsky (2021). Systematic Review and Meta-Analysis of Stay-Play-Talk Interventions for Improving Social Behaviors of Young Children. https://doi.org/10.1177/1098300720983521

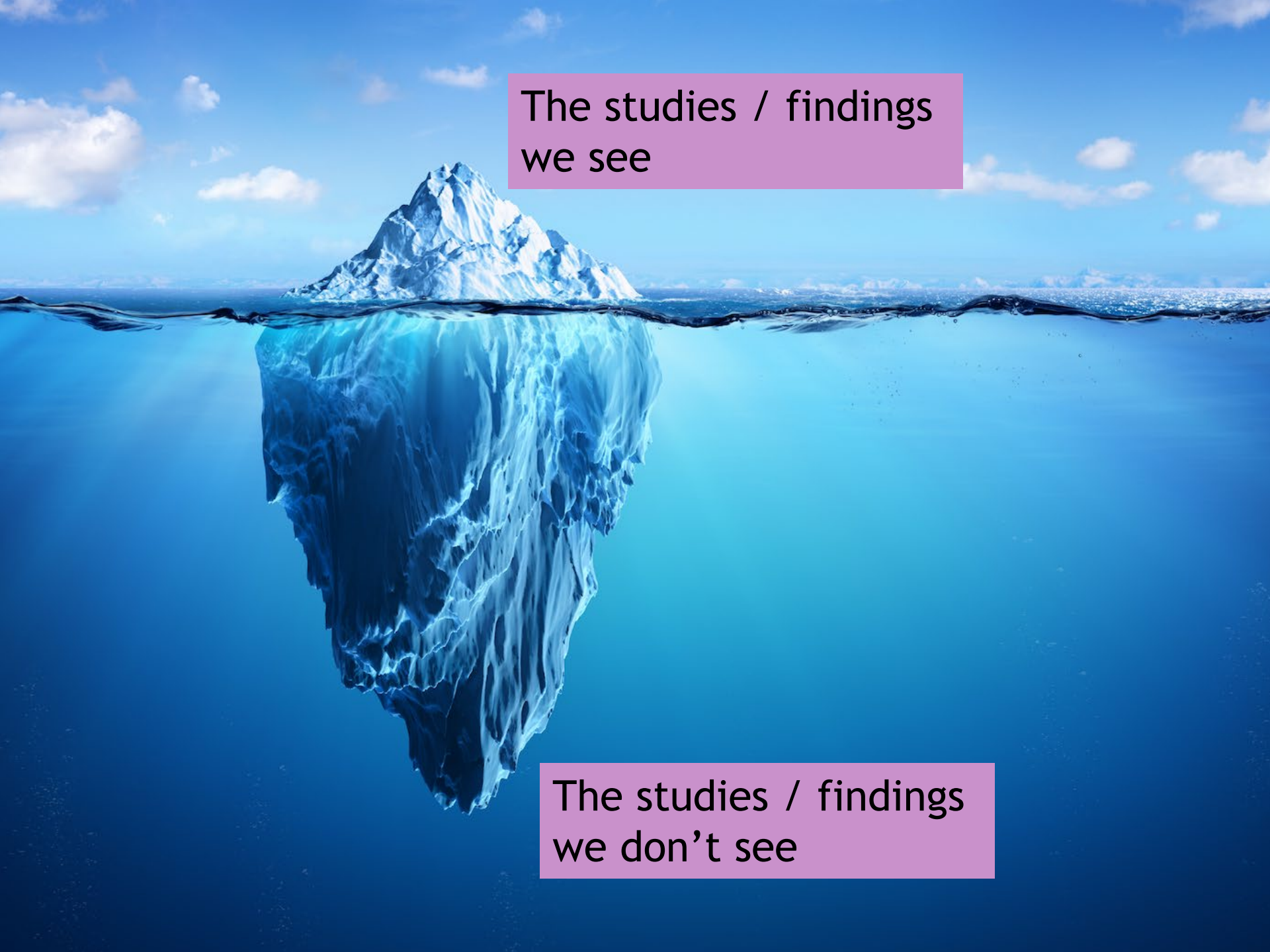
Dimension	Criteria
Participants	- 0-8 YO - Identified disability, delay, or risks, peer implementers
Intervention	- Stay-Play-Talk (SPT) intervention - Peer trained to remain in proximity and interact with “buddy”
Comparison	- Baseline
Outcomes	- Social behavior of peers and focal participants
Setting	- Free operant play activities
Design	- Experimental single-case designs

CHALLENGES

Challenge #1: Identifying all relevant evidence

- ***Publication bias***: Certain types of results may be more likely to be published, so that the published literature is not representative of the full "population" of findings.

- ***Reporting bias***: Certain types of results may be more likely to be ***reported*** (i.e., included in a research write-up), so that results included in published (or even unpublished) write-ups are not representative of the full "population" of findings.


The studies / findings we see
The studies / findings we don’t see

Publication & reporting bias in single-case research

- Good theoretical reasons to expect that publication and reporting biases affect single-case research.
  - Strong emphasis on experimental control, visually detectable functional relationships (Tincanci & Travers, 2017, 2019).

- Some empirical evidence that publication bias exists in single-case literature.
  - Sham & Smith (2014) found that findings from published studies were larger than those from unpublished dissertations in a synthesis of SCDs on pivotal response training.
  - Dowdy, Tincani, & Schneider (2020) found differences in effect size between published and unpublished studies on response interruption and redirection.
  - Single-case researchers report that they are more likely to submit/accept for publication studies with larger effects (Shadish et al., 2016).

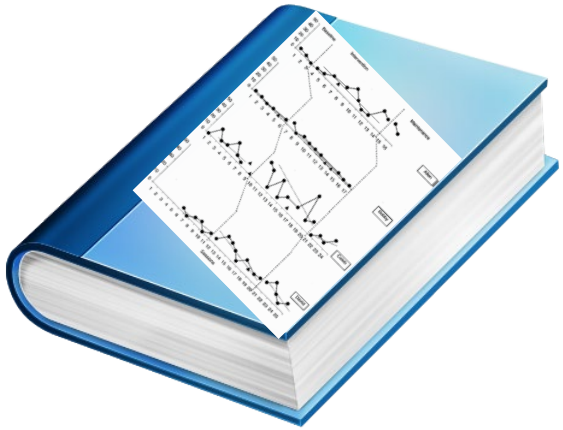
Challenge #2: Comparing findings across varied sources

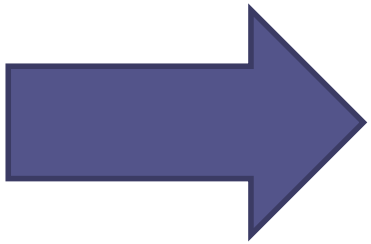
Effect size **noun**
(i-'fekt 'sīz)

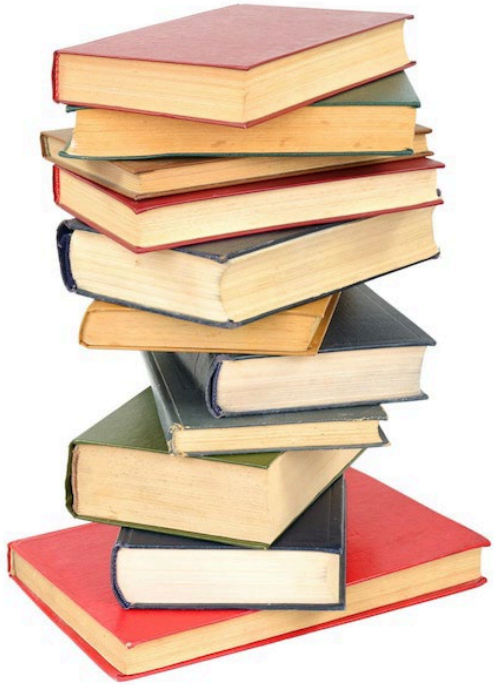
- A quantitative index describing the direction and magnitude of effects of intervention on an outcome, in a way that allows for comparison across cases and studies (Pustejovsky & Ferron, 2017).

Selecting an effect size

- The goal is to relate findings from a given study to a broader literature.
  - Effect size should describe an intervention's effects in a way that makes sense beyond the context of the original study.
  - Challenging when the set of studies has varied features (study designs, DV measurement systems)







Effect sizes for single-case research

Non-overlap measures

- Non-overlap of all pairs (Parker & Vannest, 2009)
- Tau-AB (Parker, Vannest, Davis, & Sauber 2011)
- Percentage of non-overlapping data (PND; Scruggs et al., 1987)
- Percentage exceeding the median (PEM; Ma, 2006)
- Others: PAND, RIRD, Tau-U,…

Parametric within-case measures

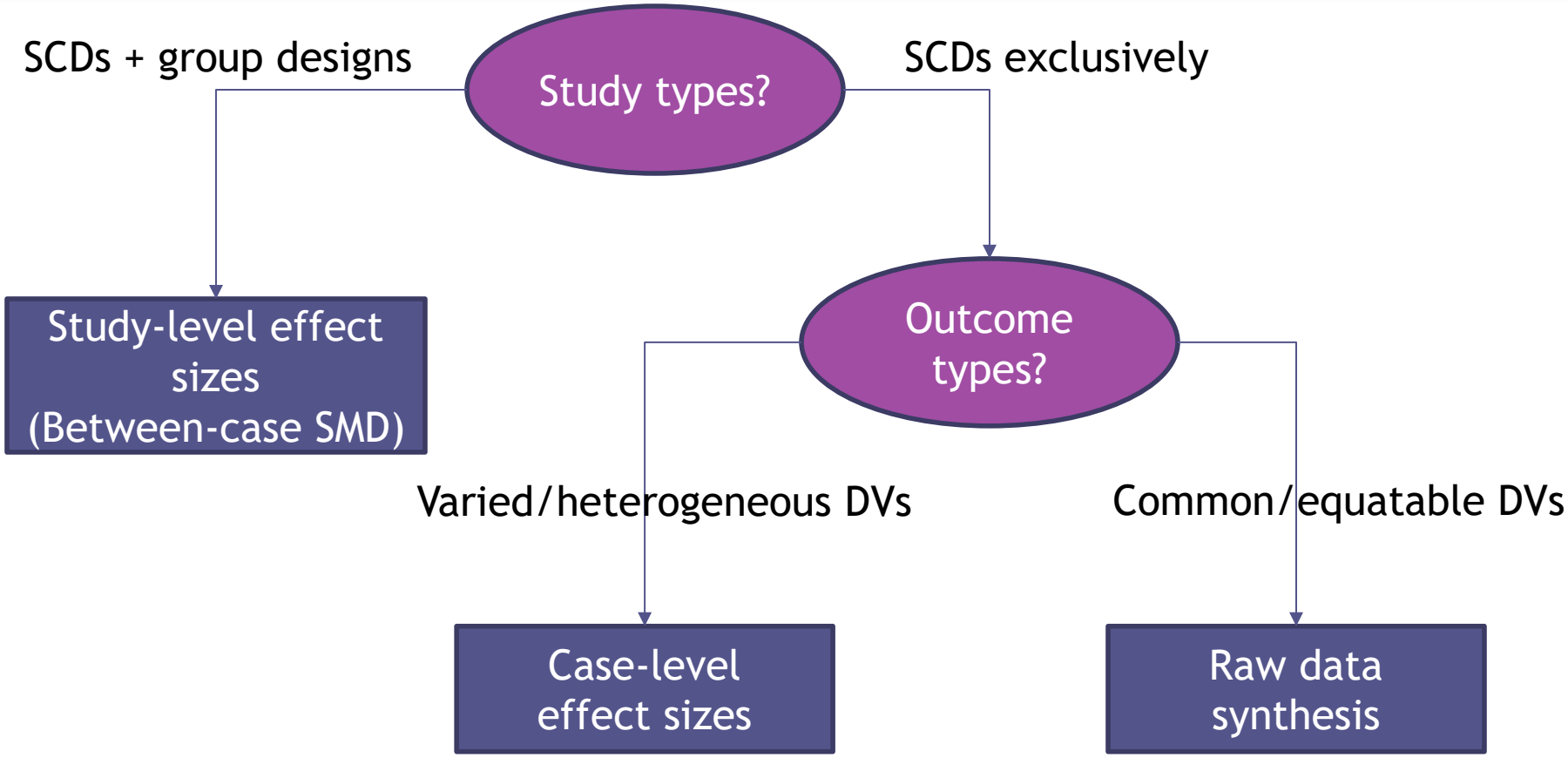
- Within-case standardized mean differences (Busk & Serlin, 1992)
- Response ratio/log-response ratio (Pustejovsky, 2018)
- Ratio of medians (Bonett & Price, 2020)
- Odds ratio / log-odds ratio (Pustejovsky, 2015)
- Percentage of Goal Obtained (Ferron, Goldstein, Olszewski, & Rohrer, 2020)

Between-case standardized mean difference

- Pustejovsky, Hedges, & Shadish (2014)
- Maggin, Swaminathan, Rogers, O'Keeffe, Sugai, & Horner (2014)
- Chen, Pustejovsky, Klingbeil, & Van Norman (2023)

Raw Data Synthesis

- Van den Noortgate & Onghena (2008)
- Moeyaert, Ugille, Ferron, Beretvas, & Van den Nortgate (2013, 2014)



Ferron, Kirby, Lippen, Pustejovsky, Chen, Grekov, & Machalicek (2023). Effect Size Estimation and Synthesis of Single-Case Designs: A Methods Guide. Institute of Education Sciences. U.S. Department of Education. Washington, DC. https://jepusto.github.io/SCD-Methods-Guide/

Challenge #3: Assessing risk of bias of primary studies

- Risk of bias "refers to the potential for study findings to systematically deviate from the truth due to methodological flaws in the design, conduct or analysis" (PRISMA 2020, p. xx).
  - cf. critical appraisal
  - cf. strength of evidence

- It is critical to assess risk of bias of primary studies in SRs to guide interpretation of the evidence

Challenge #3: Assessing risk of bias

- Selecting a tool
  - Many options, but important to understand the extent to which the tool captures risk of bias
- Integrating assessment of risk of bias into a SR
  - Inclusion criterion
  - Subgroup analysis
  - Moderator analysis

GOOD PRACTICIES

Guiding principles revisited

- Systematic
- Transparent
- Comprehensive
- Impartial (mitigation of personal biases / preconceptions)

Development of RQs and inclusion criteria

- Use of PICOS+D framework
- Collins et al. (2020): https://doi.org/10.1016/j.jsp.2020.10.002

Dimension	Criteria
Participants	- students in schools
Intervention	- researcher-manipulated independent variable involving peer reporting (whether vocal or written praise reports)
Comparison	- compared to no intervention or typical classroom contingencies
Outcomes	- academically engaged, disruptive, or social behavior
Setting	- implemented class-wide or in groups
Design	- single-case experimental design with at least three possible demonstrations of effect; three data points per phase

Searching for studies

- Searching multiple databases
- Integration of secondary search strategies
- Inclusion of gray literature
- A note on date restrictions

Making decisions about study inclusion

- Specifying procedures
  - Title/abstract screening
  - Full text review
- Training/relying on multiple coders and/or technological supports (e.g., MetaReviewer, Rayyan)
- Visual depiction

Variable coding

- Developing a codebook with operational definitions and examples/nonexamples
- Training/relying on multiple coders
- Missing data

Less settled areas…

- Effect measure selection and calculation
- Effect measure aggregation
- Exploration and explanation of heterogeneity in effect measures
- Prevention, detection, and/or correction of publication bias

Group exercise

- Suppose you and your group members are co-editors or editorial board reviewers for a journal that publishes systematic reviews of single-case research. Work together to fill out a checklist of questions that peer reviewers could use when evaluating a manuscript reporting a systematic review that includes single-case research studies. Consider some or all of the following sections for your checklist (feel free to edit the section headings as you see fit):
  - Aims and scope of the review
  - Evidence identification
  - Risk of bias assessment
  - Methodology
  - Transparency
  - Other