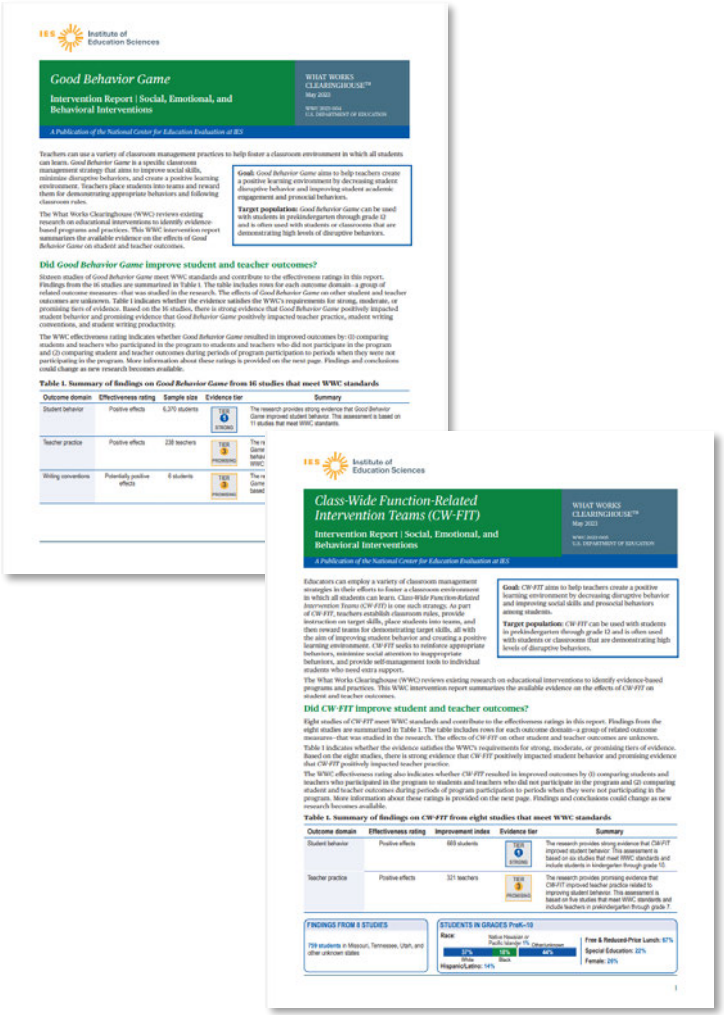
Single-Case Design Research in What Works Clearinghouse (WWC) Intervention Reports

May 2023

Why this webinar?

- **Context:** The WWC recently released two intervention reports that include findings from single-case design (SCD) studies:
  - *Good Behavior Game*
  - *Class-Wide Function-Related Intervention Teams (CW-FIT)*

- **Purpose:** This webinar will help researchers understand how the WWC reviews and synthesizes SCD research.



Road map for this webinar

- Overview of WWC products
- How SCD research expands the base of rigorous studies
- Understanding SCDs and their findings
- Explore how SCDs contribute to WWC intervention reports
- Wrap-up and where to find materials

Dr. Martha Bleeker
Principal Researcher
Mathematica

Dr. John Hitchcock
Associate Director
Westat

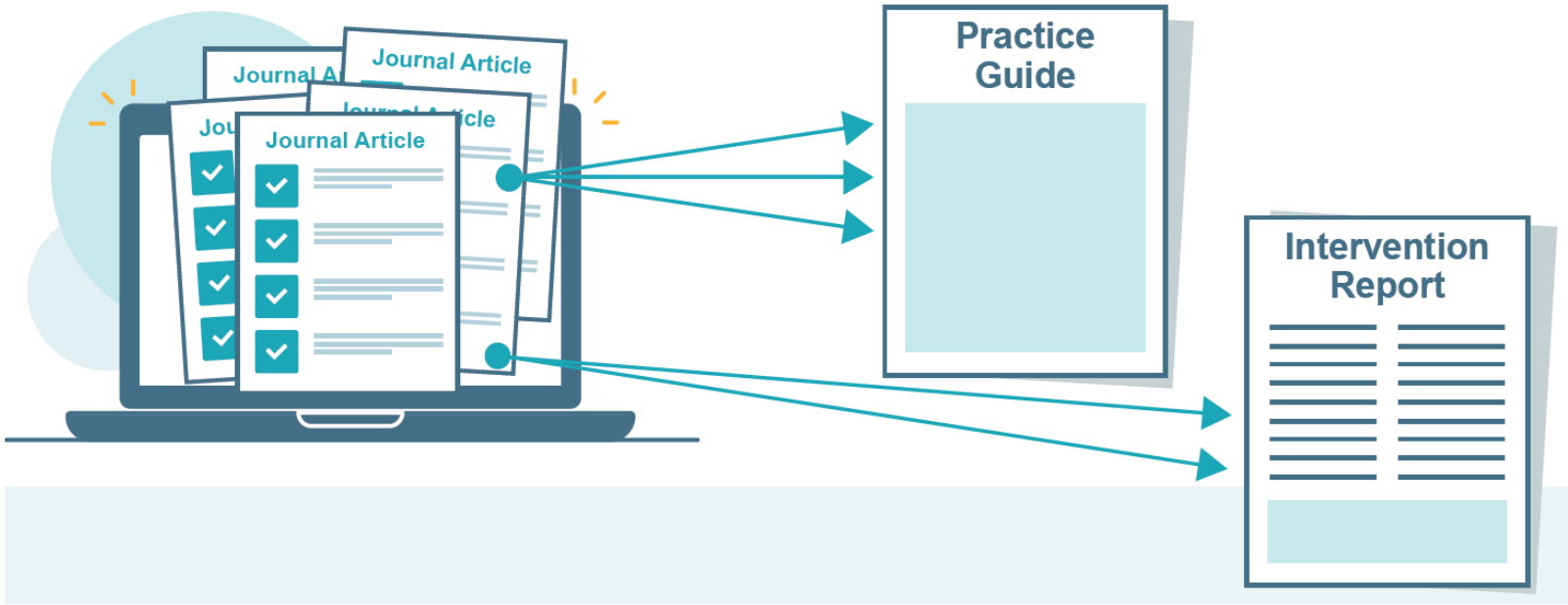
Overview of the WWC: Using evidence to improve instruction

- **Who we are:** An initiative of the Institute of Education Sciences (IES) at the U.S. Department of Education.
- **What we do:** Evaluate existing research on educational topics and summarize the evidence in accessible products.
- **Why we do it:** To provide educators and other policymakers with information to make evidence-informed decisions.
- **How we do it:** Hundreds of trained and certified reviewers rate whether studies are of sufficiently high quality and then summarize the results from the high-quality studies.

Types of WWC resources

Practice guides help educators identify and implement evidence-based strategies

Individual study reviews provide the ingredients to practice guides and intervention reports



Intervention reports help decision makers select or continue implementing a specific intervention


IES WWC What Works Clearinghouse
MENU
Search
Go
Find What Works
Search the WWC and access our Resources Page to find the information you need to make evidence-based decisions in your classrooms and schools.
Search the WWC
Enter keyword(s)
Topic (No option selected)
Grade band
Search
Practice Guides: evidence-based recommendations for educators to improve student outcomes.
Intervention Reports: summaries of findings of the highest-quality research on a given intervention or practice in education.
Reviews of Individual Studies: summaries of individual studies that have been reviewed by the WWC.
Data From Study Reviews: downloadable data from WWC study reviews.

Expanding the Evidence Base with Single-Case Design (SCD) Research

Expanding the evidence base

- Group design studies are commonly used to generate evidence of effectiveness:
  - Randomized controlled trials (RCTs)
  - Quasi-experimental designs (QEDs)
  - Regression discontinuity designs (RDDs)

  These group designs:
  - Require a comparison group that does not receive the intervention during the study
  - Can use large samples, such as all students within a class, school, or district; can be costly, especially if students are in several locations or if data are collected over a long period of time

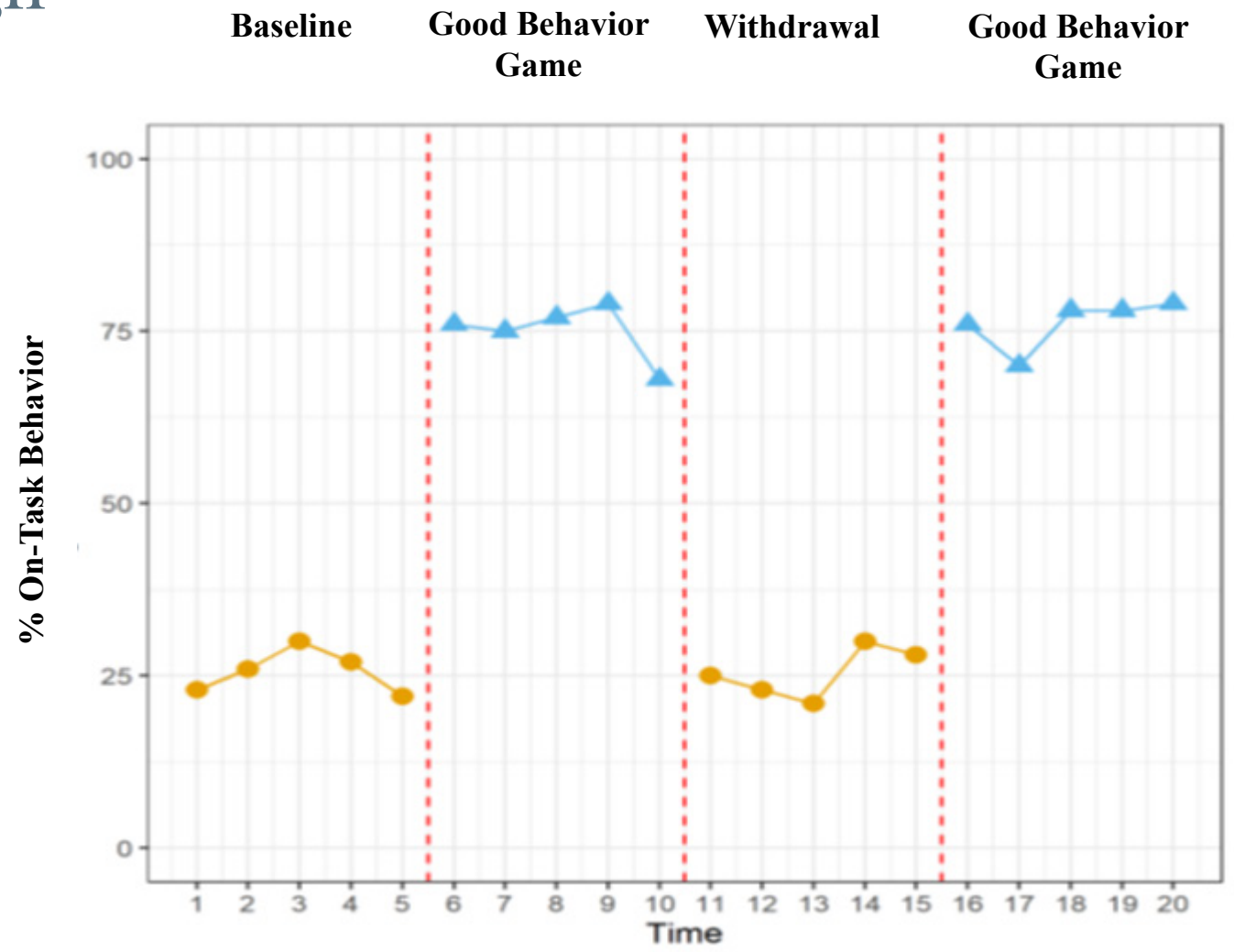
- Some populations, interventions, and outcomes are not well-suited to group design studies. For example:
  - Interventions aimed at students with specific challenges
  - Schools that cannot withhold interventions from comparison group students
  - Outcomes that are often best studied via direct observation and repeated measurements

Key features of single-case designs

- SCDs are experimental designs, where an **individual case** is the unit of intervention administration and data analysis.
- Researchers assess intervention effects in different phases, by **repeatedly observing the case before, during, and sometimes after the intervention.**
- A **case** is often a single student or group of students, such as a classroom.
- **Data are presented graphically** to illustrate the effect of the intervention by showing how outcomes change **within and across different conditions** (or phases).

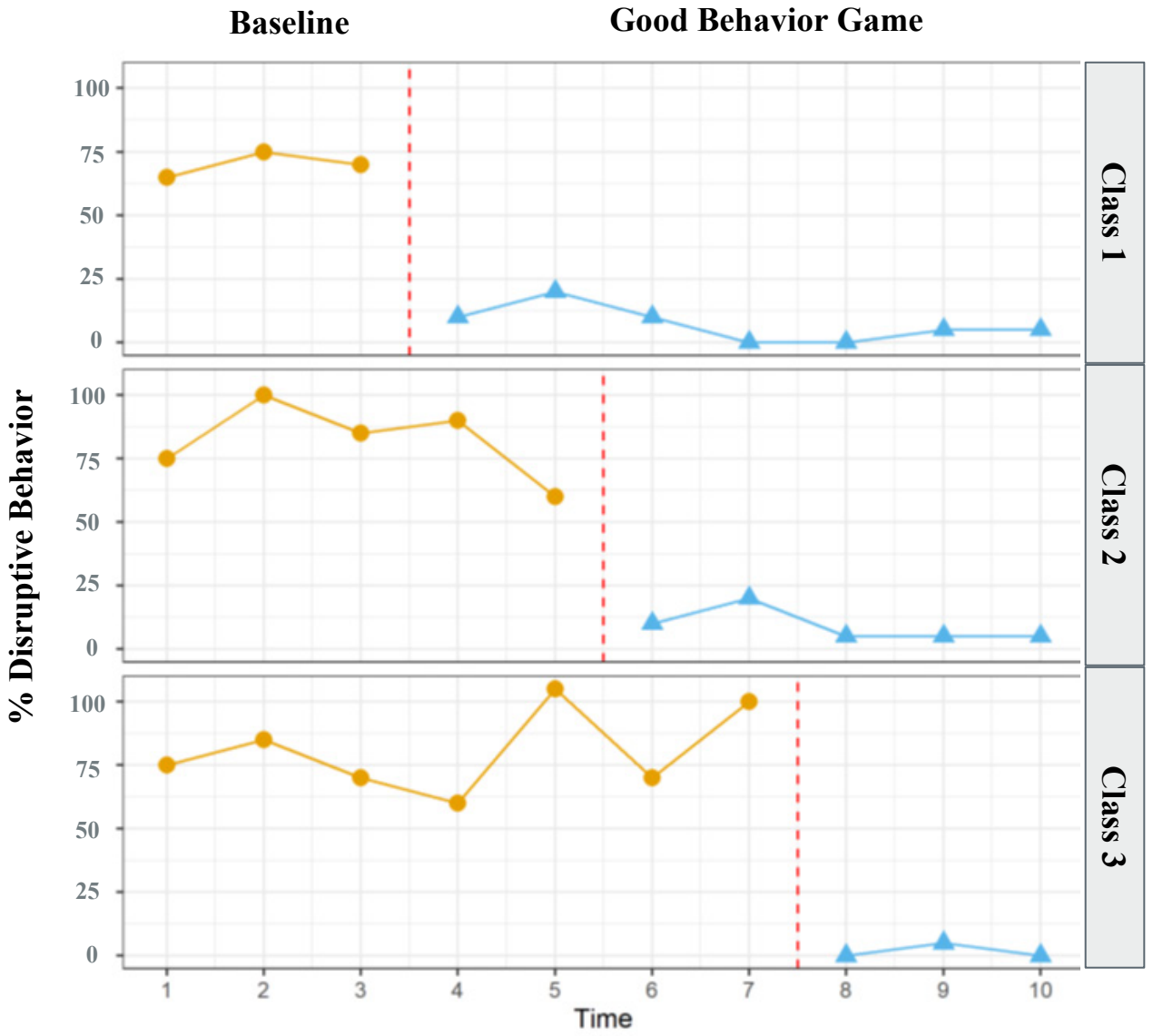
Reversal-withdrawal design

- Repeatedly introduce and remove the intervention across multiple phases
- One case provides three opportunities to demonstrate an effect at three points in time
- Appropriate when we expect effects on outcomes to reverse when intervention is removed



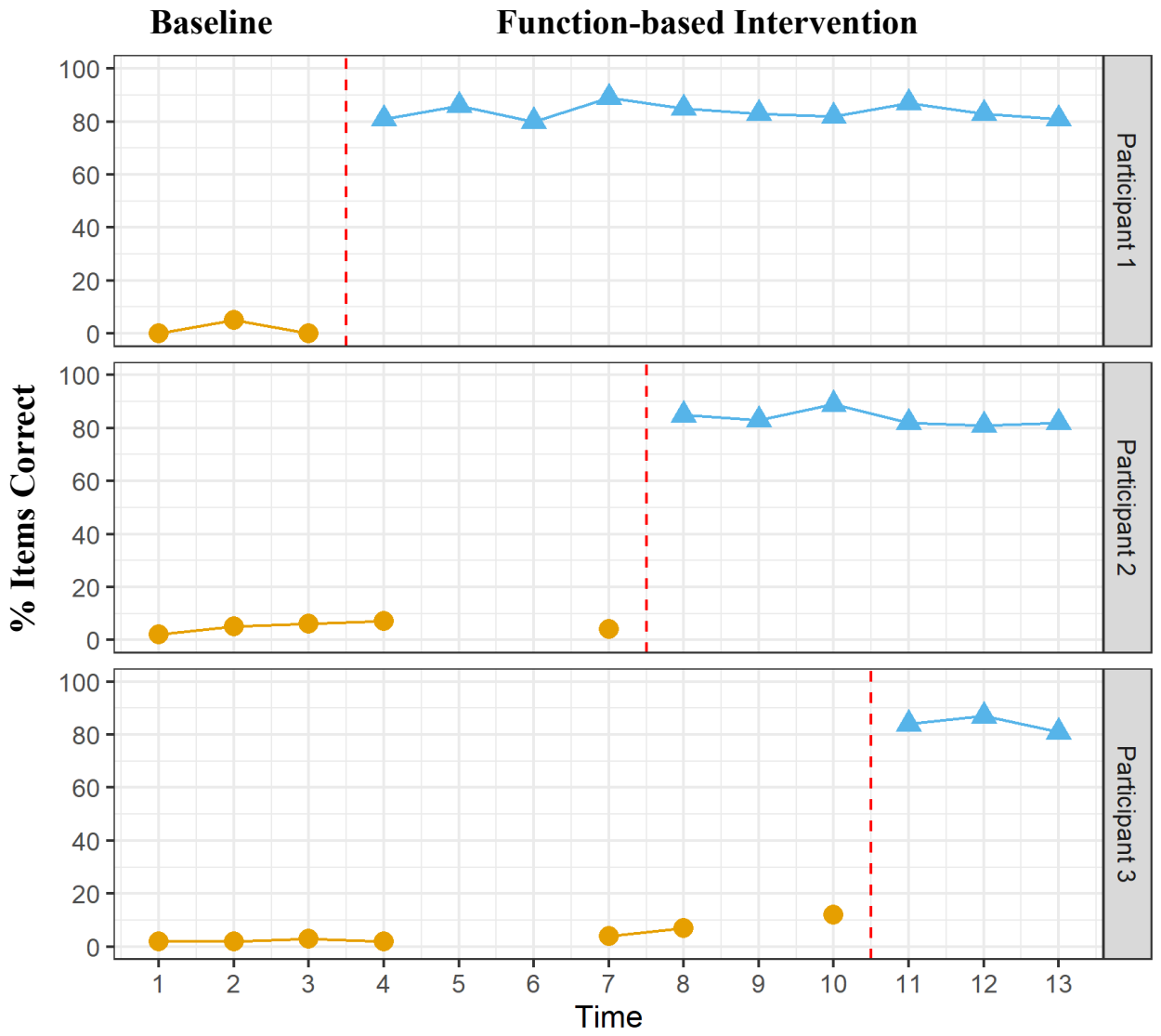
Multiple baseline design

- Stagger the introduction of the intervention across more than one case or setting
- Typically start with concurrent baseline phases; then the intervention is introduced in a staggered manner
- Each case provides an opportunity for a single demonstration of an effect of a program
- Appropriate when the intervention is expected to have a lasting effect on outcomes



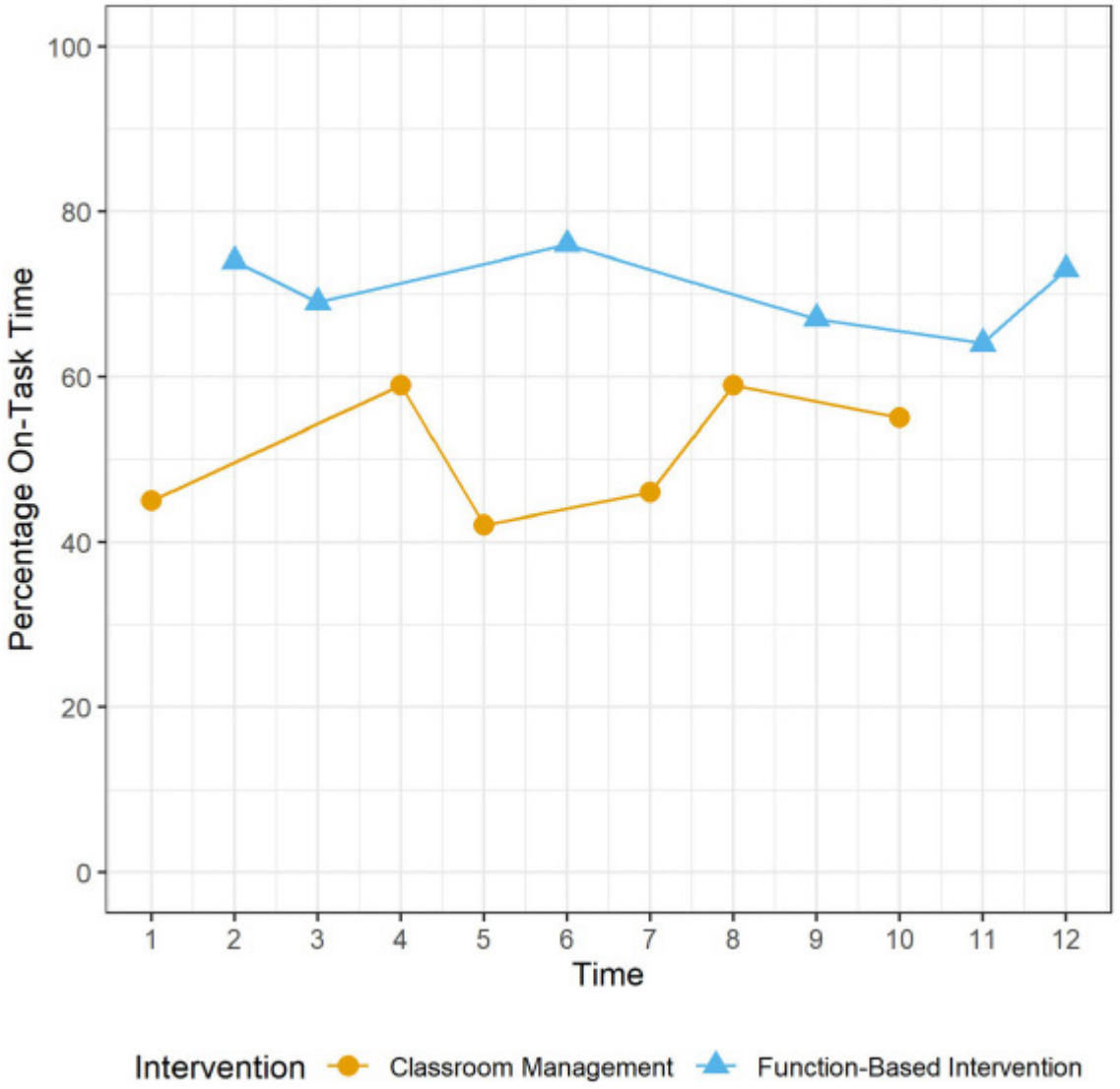
Multiple probe design

- Differentiated from a multiple baseline design by intermittent rather than continuous data collection probes.
- Missing data are a planned feature of the design.
- Particularly useful when repeated testing of a behavior or skill may be harmful or upsetting.



Alternating treatment design

- Rapid alternation between two or more interventions, with only a few observations per phase.
- Used with outcome behavior that is easily altered and likely to return to pre-intervention levels in the absence of intervention.
- May not include a baseline phase.



Key features of single-case designs

- SCD research can be used to help overcome some of the challenges with group design research.

- SCDs generally include a **small number of participants**, so can be used with smaller subsamples, such as students with high levels of disruptive behavior or students with disabilities.

- **All participants can receive the intervention** during an SCD study—there is no need to withhold services to establish a comparison group.

Requirements for SCDs to meet WWC version 4.1 standards

- SCDs must include repeated measurements across conditions.
  - At least **three** opportunities to demonstrate an effect at **three** points in time
  - Repetition improves precision of results

- SCDs can demonstrate causal effects, if researchers:
  - **Control** when the intervention is introduced (and withdrawn)
  - Carefully **monitor outcomes** in the presence and absence of the intervention

- Outcomes must be measured systematically over time by more than one assessor and meet specific inter-assessor agreement (IAA) criteria set by the WWC.
  - Percent agreement across observers must be at least **80%**
  - IAA data must be collected at least once in each phase and in at least **20%** of intervention and comparison sessions

- These requirements reduce threats to internal validity and increase the chance that observed effects were caused by the intervention.

Common reasons SCDs do not meet WWC standards (DNMS)

- Eligible outcomes do not meet requirements due to insufficient IAA data.
  - IAA data not collected at least once in each phase or in at least 20% of intervention and comparison sessions
  - Percent agreement across observers is less than 80%

  Among studies that DNMS: Represents 52% of studies in Good Behavior Game and CW-FIT

- Insufficient data to demonstrate an intervention effect.
  - Fewer than three attempts to demonstrate an intervention effect at three different points in time
  - Fewer than three data points in at least one phase of the SCD

  Among studies that DNMS: Represents 48% of studies in Good Behavior Game and CW-FIT

Reporting Results from Single-Case Designs

SCD evidence can be synthesized with findings from group design studies

- The WWC calculates and synthesizes a **design-comparable effect size** (D-CES) from SCDs with Hedges' *g* effect sizes from group design studies.
  - Conceptually similar to Hedges' *g,* the D-CES is estimated by comparing outcomes between intervention and comparison phases.
  - Accounts for trends in repeated measures of the outcome over the sessions.

- Calculating a D-CES requires three or more cases in the study.
  - May be estimated from a multiple baseline design across cases, multiple probes across cases, or a reversal-withdrawal design for three cases.
  - There is currently no procedure for computing the D-CES for other SCDs.

D-CES for outcomes measured for groups of individuals

- Some outcomes are measured for clusters or **groups of individuals**, like a classroom of students.

- The WWC will calculate a D-CES for cluster-level outcomes if they are based on **measurements of individual outcomes aggregated to the group level**.
  - For example, the percentage of students in the class exhibiting disruptive behavior

- The WWC applies a formula from the version 5.0 Procedures and Standards Handbook to transform the effect size to measure effects on students, instead of classrooms.

- The WWC cannot calculate a D-CES for cluster-level outcomes for:
  - Measurements based on scans of the cluster **without a fixed method for individually observing each student** in the cluster
  - Measurements based on **small groups of students within the cluster** instead of individual students

D-CES calculation used by WWC

- The WWC estimates D-CES using a multilevel modeling framework.
- Software products to implement D-CES estimation include the scdhlm R package or web application.
- Requires access to raw outcome data by case, observation session, and condition.
- If authors do not report raw SCD data or provide them through an author query, the WWC will extract the data from graphs using freely available software.

See p. E-10 of the version 4.1 *WWC Procedures Handbook* for an overview of D-CES calculations.

SCD influence on synthesized effects

- When synthesizing effect sizes across studies, the WWC weights each impact estimate based on its precision, so more precise estimates receive more weight.
- The precision of a D-CES is influenced by the:
  - Number of cases
  - Number of observation sessions
  - Consistency of measurements within phases
- Use of repeated measures improves precision; the D-CES can therefore contribute meaningfully to a synthesis even though sample sizes are generally much smaller than group design studies.
- SCD studies without a D-CES can still be described in WWC products but are not synthesized with other intervention findings.

WWC Intervention Reports: *Good Behavior Game* and *Class-Wide Function-Related Intervention Teams (CW-FIT)*

Next, we will:

- Describe *Good Behavior Game* and *CW-FIT*
- Summarize intervention components, implementation, and costs
- Describe the number of studies in each report and the students represented in those studies
- Briefly summarize the effectiveness of *Good Behavior Game* and *CW-FIT* across studies and outcome domains, **including findings from group design and SCD studies**

Good Behavior Game

- **Goal:** *Good Behavior Game* is a classroom management strategy that aims to help teachers create a positive learning environment by decreasing student disruptive behavior and improving student academic engagement and prosocial behaviors.
- **Target population:**
  - Students in Pre-K through grade 12
  - Often used with students or classrooms that are demonstrating high levels of disruptive behaviors
- **Implementation:** Teachers place students into teams and reward them for demonstrating appropriate behaviors and following classroom rules.
- **Training:** Teachers receive training from a developer or researcher.
  - *PAX Good Behavior Game*® and American Institutes for Research (AIR) both provide an initial 2-day training and other materials.
- **Version:** Four studies used *PAX Good Behavior Game*®, one study used the AIR version, and 11 studies did not report which approach was used.

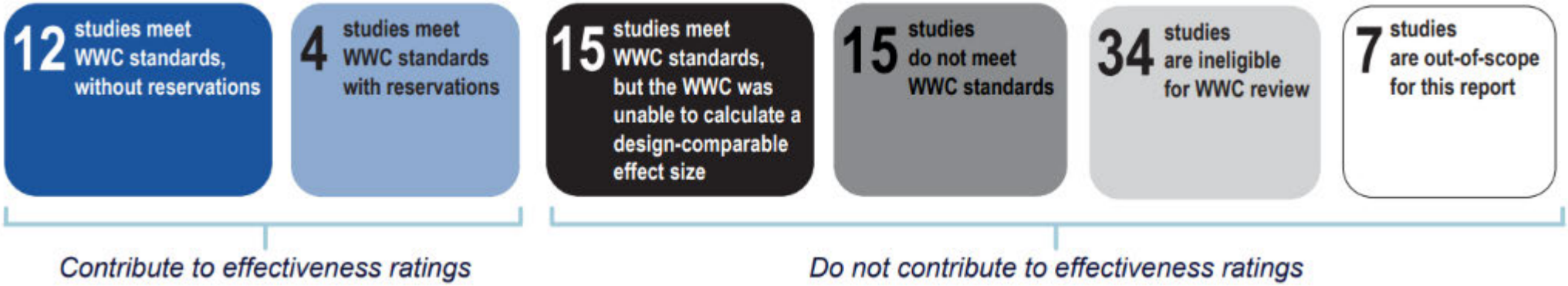
Good Behavior Game Implementation and Costs

- **Key components:**
  - Team-based game
    - Develop rules for student behavior
    - Monitor and score student behavior
    - Provide rewards to winning team(s)
  - Training for teachers
  - Ongoing coaching for teachers

- **Costs:**
  - *PAX Good Behavior Game*® ranges from **$405 per teacher** for a group training with 30 teachers to **$580 per teacher** for individual, self-paced training.
  - American Institutes for Research version is **$325 per teacher** for a group training with 40 teachers, **plus additional costs** to cover the AIR trainers' travel.

Good Behavior Game: Number of studies

- **16 of 87 studies** are eligible for review, meet WWC standards, and contribute to effectiveness ratings.
  - Includes 6 group design studies and 10 SCD studies



- **15 other SCD studies meet standards** but do not contribute to the findings:
  - 14 studies did not have at least 3 cases; 1 study used alternating treatments design.
  - The study authors' visual analysis of the findings is described in intervention report appendix, but the findings are not included in the main intervention report.

Good Behavior Game: Sample characteristics

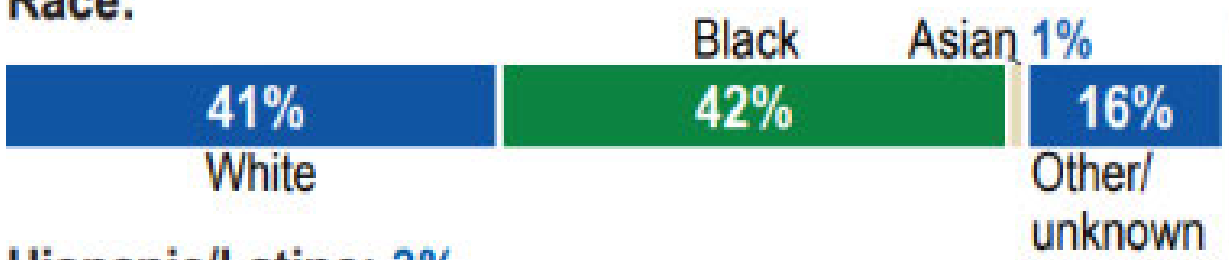
FINDINGS FROM 16 STUDIES

8,387 students in Maryland, Pennsylvania, Texas, and other unknown states; Northern Ireland; and the United Kingdom.

- Student demographic information is drawn from published studies.
- WWC asks authors for non-reported sample information, unless it could allow readers to identify sample members.

STUDENTS IN GRADES K–11

Race:



Hispanic/Latino: 3%

Free & Reduced-Price Lunch: 61%

Special Education: 15%

Female: 49%

Good Behavior Game: Findings

Table 1. Summary of findings on *Good Behavior Game* from 16 studies that meet WWC standards

Outcome domain	Effectiveness rating	Sample size	Evidence tier	Summary
Student behavior	Positive effects	6,370 students	TIER 1 STRONG	The research provides strong evidence that *Good Behavior Game* improved student behavior. This assessment is based on 11 studies that meet WWC standards.
Teacher practice	Positive effects	238 teachers	TIER 3 PROMISING	The research provides promising evidence that *Good Behavior Game* improved teacher practice related to improving student behavior. This assessment is based on three studies that meet WWC standards.
Writing conventions	Potentially positive effects	6 students	TIER 3 PROMISING	The research provides promising evidence that *Good Behavior Game* improved student writing conventions. This assessment is based on one study that meets WWC standards.
Writing productivity	Potentially positive effects	6 students	TIER 3 PROMISING	The research provides promising evidence that *Good Behavior Game* improved student writing productivity. This assessment is based on one study that meets WWC standards.
Literacy achievement	Uncertain effects	3,453 students	NO TIER ASSIGNED	The research does not support claims that *Good Behavior Game* improved student literacy achievement. This assessment is based on two studies that meet WWC standards.
Mathematics achievement	Uncertain effects	703 students	NO TIER ASSIGNED	The research does not support claims that *Good Behavior Game* improved student mathematics achievement. This assessment is based on one study that meets WWC standards.
Intrapersonal competencies	Uncertain effects	3,857 students	NO TIER ASSIGNED	The research does not support claims that *Good Behavior Game* improved student intrapersonal competencies. This assessment is based on two studies that meet WWC standards.
School climate	Uncertain effects	73 after-school programs	NO TIER ASSIGNED	The research does not support claims that *Good Behavior Game* improved school climate. This assessment is based on one study that meets WWC standards.

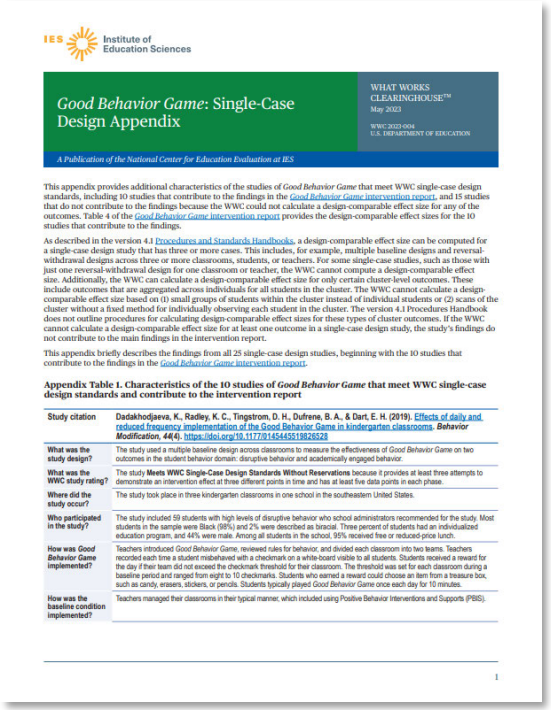
Good Behavior Game: Findings *(continued)*

Table 4b. Findings by outcome domain from three studies of *Good Behavior Game* that meet WWC standards: Teacher practice outcomes

Outcome	Timing of measurement	Study sample	Number of sites	Mean (standard deviation)		Findings		
				Intervention group	Comparison group	Effect size	Improvement index	Statistically significant (*p*-value)
Teacher practice outcome domain								
Ohio State Teachers' Sense of Efficacy Scale: Classroom Management Subscale	End of intervention (Humphrey et al., 2018)	230 teachers in grades 2 and 3	77 schools in the United Kingdom	8.18	8.09	0.06	+2	No (p=0.35)
Behavior-specific teacher praise, based on researcher observation of teachers	During SCD intervention sessions (Lynne et al., 2017)	3 teachers in grades 1 and 4	1 school in southwestern U.S.	--	--	0.67	+25	No (p=0.12)
General teacher praise, based on researcher observation of teachers	During SCD intervention sessions (Lynne et al., 2017)	3 teachers in grades 1 and 4	1 school in southwestern U.S.	--	--	-0.10	-4	No (p=0.66)
Instructor correction for problem social behavior, based on researcher observation of teachers	During SCD intervention sessions (Rodriguez, 2010)	5 instructional assistants in grade K	1 school in the Pacific Northwest U.S.	--	--	0.75	+27	Yes (p<0.01)
Instructor praise for social behavior, based on researcher observation of teachers	During SCD intervention sessions (Rodriguez, 2010)	5 instructional assistants in grade K	1 school in the Pacific Northwest U.S.	--	--	1.49	+43	Yes (p<0.01)
Summary for teacher practice: positive effects						**0.35**	**+14**	**Yes (p<0.01)**

Good Behavior Game: Findings *(continued)*

- An appendix provides additional information about the 15 other SCD studies that meet WWC standards but do not contribute to the findings, including:
  - Study and sample characteristics
  - Authors' characterization of findings
- The appendix also provides information about that 10 SCD studies that *did contribute* to findings, including:
  - Study and sample characteristics
  - Authors' characterization of findings for outcomes where a D-CES could not be calculated
- WWC users can also access public extract data from each study.



IES Institute of Education Sciences

Good Behavior Game: Single-Case Design Appendix

WHAT WORKS CLEARINGHOUSE™ May 2023

WWC 2023-004 U.S. DEPARTMENT OF EDUCATION

A Publication of the National Center for Education Evaluation at IES

This appendix provides additional characteristics of the studies of *Good Behavior Game* that meet WWC single-case design standards, including 10 studies that contribute to the findings in the *Good Behavior Game* intervention report, and 15 studies that do not contribute to the findings because the WWC could not calculate a design-comparable effect size for any of the outcomes. Table 4 of the *Good Behavior Game* intervention report provides the design-comparable effect sizes for the 10 studies that contribute to the findings.

As described in the version 4.1 Procedures and Standards Handbooks, a design-comparable effect size can be computed for a single-case design study that has three or more cases. This includes, for example, multiple baseline designs and reversal-withdrawal designs across three or more classrooms, students, or teachers. For some single-case studies, such as those with just one reversal-withdrawal design for one classroom or teacher, the WWC cannot compute a design-comparable effect size. Additionally, the WWC can calculate a design-comparable effect size for only certain cluster-level outcomes. These include outcomes that are aggregated across individuals for all students in the cluster. The WWC cannot calculate a design-comparable effect size based on (1) small groups of students within the cluster instead of individual students or (2) scans of the cluster without a fixed method for individually observing each student in the cluster. The version 4.1 Procedures Handbook does not outline procedures for calculating design-comparable effect sizes for these types of cluster outcomes. If the WWC cannot calculate a design-comparable effect size for at least one outcome in a single-case design study, the study's findings do not contribute to the main findings in the intervention report.

This appendix briefly describes the findings from all 25 single-case design studies, beginning with the 10 studies that contribute to the findings in the *Good Behavior Game* intervention report.

Appendix Table 1. Characteristics of the 10 studies of *Good Behavior Game* that meet WWC single-case design standards and contribute to the intervention report

Study citation	Dadakhodjaeva, K., Radley, K. C., Tingstrom, D. H., Dufrene, B. A., & Dart, E. H. (2019). Effects of daily and reduced frequency implementation of the Good Behavior Game in kindergarten classrooms. *Behavior Modification, 44*(4). https://doi.org/10.1177/0145445519826528
What was the study design?	The study used a multiple baseline design across classrooms to measure the effectiveness of *Good Behavior Game* on two outcomes in the student behavior domain: disruptive behavior and academically engaged behavior.
What was the WWC study rating?	The study **Meets WWC Single-Case Design Standards Without Reservations** because it provides at least three attempts to demonstrate an intervention effect at three different points in time and has at least five data points in each phase.
Where did the study occur?	The study took place in three kindergarten classrooms in one school in the southeastern United States.
Who participated in the study?	The study included 59 students with high levels of disruptive behavior who school administrators recommended for the study. Most students in the sample were Black (98%) and 2% were described as biracial. Three percent of students had an individualized education program, and 44% were male. Among all students in the school, 95% received free or reduced-price lunch.
How was *Good Behavior Game* implemented?	Teachers introduced *Good Behavior Game*, reviewed rules for behavior, and divided each classroom into two teams. Teachers recorded each time a student misbehaved with a checkmark on a white board visible to all students. Students received a reward for the day if their team did not exceed the checkmark threshold for their classroom. The threshold was set for each classroom during a baseline period and ranged from eight to 10 checkmarks. Students who earned a reward could choose an item from a treasure box, such as candy, erasers, stickers, or pencils. Students typically played *Good Behavior Game* once each day for 10 minutes.
How was the baseline condition implemented?	Teachers managed their classrooms in their typical manner, which included using Positive Behavior Interventions and Supports (PBIS).

1

Class-Wide Function-Related Intervention Teams (CW-FIT)

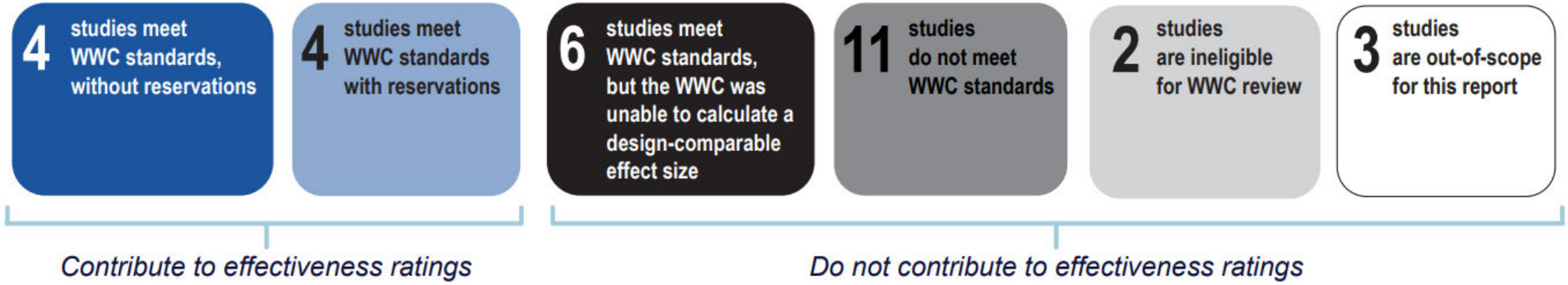
- **Goal:** *CW-FIT* aims to help teachers create a positive learning environment by decreasing disruptive behavior among students and improving social skills and prosocial behaviors.
- **Target population:** *CW-FIT* can be used with students in Pre-K through grade 12 and is often used with students or classrooms that are demonstrating high levels of disruptive behaviors.
- **Implementation:** Teachers establish classroom rules, provide instruction on target social skills, place students into teams, and reward them for demonstrating the skills.
- **Training:** Teachers can receive training from another staff member or from a *CW-FIT* researcher or developer, or they can train themselves.
  - Materials are available for free on the developer's website.

CW-FIT Implementation and Costs

- **Key components:**
  - Teach students behavioral skills
  - CW-FIT game sessions
  - Praise students
  - Additional, targeted supports for students
  - Training and support for teachers
- **Costs:**
  - **Teacher training:** In all eight studies, teachers were trained using materials available for free on the CW-FIT developer's website or provided by the study authors.
  - **Coaches:** In six studies, teachers received additional coaching. Districts can purchase training packages for coaches, starting at $6,000 to train two district coaches who can support 20 teachers.

CW-FIT: Number of studies

- **8 of 30 studies** are eligible for review, meet WWC standards, and contribute to effectiveness ratings.
  - Includes 2 group design studies and 6 SCD studies



- **6 other SCD studies** do not contribute to the findings:
  - 5 studies had fewer than 3 cases; 1 study focused on a cluster-level outcome based on small groups of students.
  - Authors' visual analysis of the findings is described in appendix.

CW-FIT: Sample characteristics

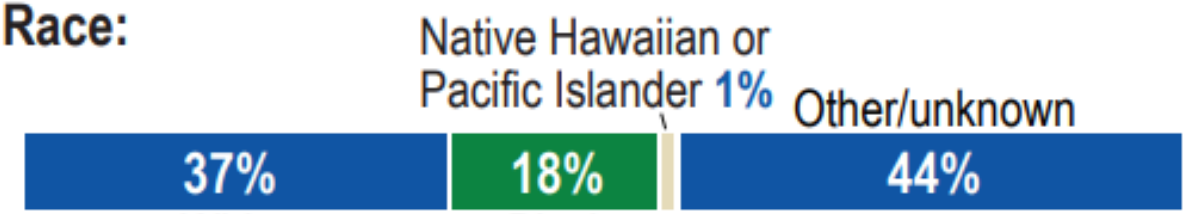
FINDINGS FROM 8 STUDIES

759 students in Missouri, Tennessee, Utah, and other unknown states

- Student demographic information is drawn from published studies.
- WWC asks authors for non-reported sample information, unless it could allow readers to identify sample members.

STUDENTS IN GRADES PreK–10

Race:



Hispanic/Latino: 14%

Free & Reduced-Price Lunch: 67%

Special Education: 22%

Female: 26%

CW-FIT: Findings

Table 1. Summary of findings on *CW-FIT* from eight studies that meet WWC standards

Outcome domain	Effectiveness rating	Improvement index	Evidence tier	Summary
Student behavior	Positive effects	669 students	TIER 1 STRONG	The research provides strong evidence that *CW-FIT* improved student behavior. This assessment is based on six studies that meet WWC standards and include students in kindergarten through grade 10.
Teacher practice	Positive effects	321 teachers	TIER 3 PROMISING	The research provides promising evidence that *CW-FIT* improved teacher practice related to improving student behavior. This assessment is based on five studies that meet WWC standards and include teachers in prekindergarten through grade 7.

- An appendix provides additional information about the 6 SCD studies that *contribute* to the findings, as well as the 6 SCD studies that *did not contribute* to findings, including:
  - Study and sample characteristics
  - Authors' characterization of findings without a D-CES

Key takeaways

- SCDs can be used with small sample sizes and do not require withholding the intervention from a comparison group.
- SCDs can demonstrate causal effects, if researchers:
  - Control the introduction and withdrawal of the intervention
  - Use repeated measures to monitor outcomes in the presence and absence of the intervention
- The WWC provides:
  - Standards to evaluate rigor of SCDs
  - Guidelines for synthesizing SCD findings with findings from group design studies
- WWC intervention reports draw on SCD results to complete the picture of an intervention's effects.

Resources

- *Good Behavior Game* Intervention Report
- *CW-FIT* Intervention Report

WWC resources used for the *Good Behavior Game* and *CW-FIT* intervention reviews:

- WWC Version 4.1 Standards Handbook
- WWC Version 4.1 Procedures Handbook
- WWC Study Review Protocol, Version 4.1

WWC Version 5.0 resources are now available:

- WWC Version 5 Procedures and Standards Handbook
- Version 5.0 Group Design online training
- Version 5.0 SCD online training

Thank you for your interest in this webinar!

- If you have any questions about the WWC, intervention reports, or any other WWC products, you can contact the WWC Help Desk (https://ies.ed.gov/ncee/wwc/help) or email Contact.WWC@ed.gov.