

Summary of Changes in the *WWC Procedures and Standards Handbook, Version 5.0*

WWC 2022008
U.S. DEPARTMENT OF EDUCATION

A Publication of the National Center for Education Evaluation at IES

The mission of the [What Works Clearinghouse](#) (WWC) is to be a central and trusted source of scientific evidence for what works in education. The WWC reviews relevant research, identifies well-designed and well-implemented impact studies, summarizes the findings from those studies, and disseminates them to the public. The goal of the WWC is to help educators, administrators, families, researchers, and policymakers make evidence-based decisions.

In August 2022, the WWC released [version 5.0 of the *WWC Procedures and Standards Handbook*](#). This summary document describes major changes between version 4.1 and version 5.0 of WWC procedures and standards, and it provides the WWC's rationale for these changes.

The Institute of Education Sciences, which oversees the WWC, consulted with the WWC's Statistical, Technical, and Analysis Team (STAT)—which includes outside methodological consultants, as well as key staff from different WWC contractors—on updates in the [WWC *Procedures and Standards Handbook, Version 5.0*](#).

The *WWC Procedures and Standards Handbook, Version 5.0*, is a single document.

Version 5.0 of the *Handbook* replaces the two documents used since October 2020: the [What Works Clearinghouse *Procedures Handbook, Version 4.1*](#), and the [What Works Clearinghouse *Standards Handbook, Version 4.1*](#). The WWC chose to combine procedures and standards into one document to improve the *Handbook's* usability to education practitioners, WWC reviewers, and others who rely on the WWC to assess the quality of education research.

Experts consulted for version 5.0 updates to WWC procedures and standards

WWC STAT and key contractor staff: Jack Buckley, Molly Cain, Sarah Caverly, John Ferron, Alicia Garcia, Michael S. Garet, Russell Gersten, Natalya Gnedko-Berry, Ben B. Hansen, Fran Harmon, Larry V. Hedges, Daniel Hubbard, Wendy Machalicek, Rebecca Maynard, David Miller, Hiren Nisar, Terri Pigott, Joshua Polanin, Allan Porowski, James E. Pustejovsky, Jordan Rickles, David Rindskopf, Sarah Sahni, Lisa Shimmel, Jessaca Spybrook, Daniel M. Swan, Emily Tanner-Smith, Joe Taylor, Elizabeth Tipton, Jeffrey C. Valentine, Elias Walsh, Ryan Williams, and Vivian Wong.

Key Institute of Education Sciences staff: Jonathan Jacobson, Erin Pollard, and Betsy Wolf.

WHAT'S NEW IN VERSION 5.0 OF THE WWC PROCEDURES AND STANDARDS HANDBOOK?

Major revisions in [version 5.0 of the WWC Procedures and Standards Handbook](#) are grouped below into changes to the organization of the *Handbook*, followed by changes to the procedures and standards.

Changes to the organization of the *Handbook*

The WWC made several changes to the organization of the [WWC Procedures and Standards Handbook, Version 5.0](#).

1. The WWC added a **new chapter** to the *Handbook*: Chapter I, Overview of the What Works Clearinghouse and Its Procedures and Standards. This new chapter provides a succinct summary of the WWC and is intended for a broad audience of practitioners, policymakers, and researchers.
2. The WWC modified the *Handbook's* organization to **align the flow of content with the study review process**. The WWC also **expanded the technical appendices** in version 5.0 of the *Handbook* to separate technical information previously contained in the main body of the [WWC Procedures Handbook, Version 4.1](#) and the [WWC Standards Handbook, Version 4.1](#). The WWC made these changes to improve the *Handbook's* usability for WWC reviewers.

Changes to procedures and standards

The WWC made several changes to the procedures and standards in the [WWC Procedures and Standards Handbook, Version 5.0](#).

3. Under previous versions, the handbooks articulated the standards for study reviews, but topic area review teams had the ability to customize certain aspects of the standards. Under version 5.0, **the WWC no longer allows for topic-specific customization of the standards**. The WWC shifted to a uniform application of the standards because allowing topic area customization of some standards resulted in the same study having multiple and sometimes different WWC ratings, creating inconsistency and confusion. In some instances, the WWC standards include different options for application, with the choice dependent on the circumstances of individual studies. Teams reviewing individual studies will decide on the most appropriate option for each study. For example, determining joiner risk is now a review team decision based on the circumstances of individual studies but is not a topic area decision.
4. Under version 5.0, **all WWC study reviews will be conducted according to the WWC Procedures and Standards Handbook and complemented by the [Study Review Protocol](#)**, including individual study reviews and reviews of studies included in evidence synthesis products, such as practice guides and intervention reports. The [Study Review Protocol](#) articulates information to supplement the *Handbook*, such as how outcome measures should be grouped into outcome domains. All studies reviewed individually and as part of evidence synthesis products will be reviewed using the [Study Review Protocol](#) to increase review consistency, transparency, and efficiency. Topic area synthesis protocols will continue to be used to provide criteria for the

literature search; guidance on how to identify and prioritize relevant studies for review and inclusion in evidence synthesis products; and guidance on intervention, sample, and outcome eligibility criteria for the synthesis.

5. **The WWC aligned its effectiveness ratings with [U.S. Department of Education evidence definitions](#)** for individual studies and synthesis products. The WWC modified effectiveness ratings to include evidence definitions to streamline the identification of effective interventions.
6. Under version 4.1, the study or intervention effectiveness rating was the highest rating obtained from any main finding, with *main findings* usually being findings for the full study sample at the end of intervention. Under version 5.0, **the WWC determines effectiveness ratings at the outcome domain level**. If a study has multiple main findings in the same domain, the WWC creates a composite finding and reports the effectiveness rating for the domain-level composite. The WWC shifted to domain-level composites because a synthesis of multiple findings in the same domain will tend to provide a better representation of the underlying construct than will any single measurement. In addition, this change simplifies WWC procedures by eliminating the need for multiple comparison corrections.
7. A new procedure under version 5.0 allows the WWC's effectiveness ratings for individual study reviews to be based on **outcome measures that are independent** of intervention developers and study authors. This procedure reflects the WWC's concern that potentially meaningful differences in effect sizes can be obtained from measures created by intervention developers or study authors, and that these measures may not be as informative to policymakers and practitioners as independent measures. Therefore, in outcome domains that have a relatively plentiful number of recognized, widely accepted, and independent measures, the WWC review will focus on those measures for reporting on an intervention's effectiveness. The [Study Review Protocol](#) will identify outcome domains for which the WWC will use independent measures to assess effectiveness. Studies that use nonindependent measures can still meet WWC standards. Nonindependent measures can contribute to a cross-study synthesis (for example, intervention reports, practice guide) when the need for nonindependent measures is documented in the topic area review protocol, but nonindependent measures will not contribute to effectiveness ratings in individual study reviews.
8. Under version 5.0, when a cross-study synthesis includes findings rated *Meets WWC Standards Without Reservations* and *Meets WWC Standards With Reservations* and the sample size is sufficiently large, the **WWC will attempt to ensure that a majority of the meta-analytic weight is based on findings rated *Meets WWC Standards Without Reservations***. The rationale for this change is to ensure that findings from the most rigorously designed studies receive the most weight in the synthesis.
9. Under version 5.0, individual-level and cluster-level **high attrition randomized controlled trials (RCTs) and high attrition regression discontinuity design (RDD) studies no longer need to demonstrate baseline equivalence** to be rated *Meets WWC Standards With Reservations* when attrition bias is assessed using the optimistic boundary. Study authors only need to use an

acceptable adjustment strategy in the impact analysis. The WWC allows this flexibility because while attrition can undermine the validity of an estimated intervention effect, strong control over the assignment mechanism (through randomization in RCTs or a forcing variable cutoff in RDDs) often provides a reasonable basis for statistical procedures that attempt to adjust for the potentially biasing effects of attrition. This change also allowed the WWC to bring RCT and RDD reviews into closer alignment than in previous versions of the standards.

10. Under previous versions of the standards, review protocols determined the choice between the optimistic and cautious attrition boundaries. Under version 5.0, **teams conducting WWC reviews are responsible for determining whether to use an optimistic or a cautious attrition boundary** for a specific review and for documenting their reasoning based on the principles described in the Handbooks. If review teams find that they cannot defensibly choose between the optimistic and cautious boundaries, then they should use the cautious attrition boundary because when there is doubt, a more cautious approach is warranted. The WWC standards allow for review teams' choice of the attrition boundary because the applicability of optimistic or cautious attrition assumptions depends on the circumstances of individual studies, which review teams are best positioned to evaluate against the WWCs' guidance on selecting the attrition boundary.
11. The WWC has **removed procedures for WWC-applied difference-in-difference adjustments**, which had previously allowed the WWC to use reported baseline information to adjust effect sizes based on unadjusted outcome statistics. If a study requires baseline adjustment to meet WWC standards, then the study authors must be the ones to apply any required adjustment, not the WWC. The WWC will continue to report effect sizes based on unadjusted statistics if adjusted statistics are unavailable and adjustment for baseline differences was not required, such as for low-attrition RCTs. The change to remove WWC-applied difference-in-differences adjustments aligns with an overall principle in version 5.0 of the WWC *Handbook* of greater transparency by increasing correspondence between the effects reported by the study authors and the WWC.
12. **The WWC no longer considers bundled—or combined—interventions a confounding factor in reviews of individual studies** because a bundled intervention can produce a valid impact estimate for the “package” of interventions, provided they are eligible for WWC review. For a topic area synthesis, a bundled intervention will remain problematic if any of the bundled interventions do not meet the definition of eligibility as articulated in the topic area synthesis protocol. Under this circumstance, a bundled intervention may be excluded from a synthesis product.
13. **The WWC now classifies cluster RCTs as having either a low or high risk of bias due to compositional change from joiners** (individuals who enter intervention or comparison clusters after the clusters' assignments to conditions is known outside of the study team). The change eliminates a prior distinction about late versus early joiners, and provides explicit guidance about when including joiners in the analytic sample should not affect the study's research rating. Review teams will characterize the risk of bias due to joiners based on three factors: (a) the unit of assignment, (b) the unit of measurement, and (c) the potential for the intervention to affect joining. Review teams will use a similar set of factors to characterize the risk of bias due to leavers, which guides the choice of the attrition boundary and an acceptable reference sample for determining

individual-level attrition. The WWC made these changes to simplify the cluster-level assignment standards and decrease ambiguity in applying them.

14. Under version 5.0, single-case design studies (SCDs) that use multiple baseline/multiple probe, treatment reversal, and changing criterion designs need to have **at least six data points in the initial baseline phases** for their findings to be eligible for the rating *Meets WWC Standards Without Reservations*. Previous versions of the standards required at least five data points per phase for designs to be eligible for the rating *Meets WWC Standards Without Reservations*. This change was intended to ensure there is sufficient opportunity to understand the initial pattern of responding in these designs.
15. Version 5.0 introduces **an exception for minimum data point requirements** for SCDs that use multiple baseline/multiple probe, treatment reversal, and changing criterion designs. Any phases with three or more data points and zero within-phase variance, including the initial baseline phase, are considered to have sufficient data points to be eligible for the rating *Meets WWC Standards Without Reservations*. The WWC made this change because additional data points would likely not improve a study's design under these circumstances.
16. Under version 5.0, the WWC modified **the interobserver agreement requirements for SCDs to apply to all data** in the study rather than to specific conditions as was the case under the previous version of WWC procedures and standards. The WWC made this change to better align the standards with practice in high-quality SCDs.
17. The WWC **added a new "limit risk of bias" step to the review process** for multiple baseline/multiple probe, treatment reversal, and changing criterion SCDs that are eligible for the rating *Meets WWC Standards Without Reservations*. This process uses the nonoverlap of all pairs as a decision rule that is intended to be analogous to some of the visual-analytic judgments that are used to assess the internal validity of SCDs.
18. The WWC provided **updated guidance on how to rate SCDs** with features from multiple design types and SCDs with more cases and/or phases than the minimum required to meet WWC standards. Under version 5.0, a study is typically eligible to receive the highest rating that any subset of cases or phases is eligible to receive. The WWC made this change to ensure that SCDs that include information above and beyond what is required by the standards are not penalized for reporting more data than studies that report the minimum data required, and to allow study authors more flexibility to design studies using a combination of design features. This change also brings SCD study ratings into closer alignment with group design study ratings.