

Questions and answers about the *WWC Procedures and Standards Handbook, Version 5.0*

The purpose of this document is to provide answers to questions submitted before and during What Works Clearinghouse (WWC) webinars hosted in March 2022 and August 2022 to publicize the public comment period and subsequent release of the [WWC Procedures and Standards Handbook, Version 5.0](#). We combined questions and answers (Q&A) into a single document because most questions submitted before or addressed during both webinars were similar or duplicative.

This document is meant to serve as a companion to the webinar slide decks. We combined similar questions and rephrased others for clarity while preserving the meaning of the original questions. If additional questions arise, please contact the WWC Help Desk at <https://ies.ed.gov/ncee/wwc/help>.

Questions about WWC procedures

1. Will studies that were previously reviewed need to be rereviewed under the new standards?

Studies that were previously reviewed by the WWC are not automatically slated for rereview when a new version of the standards is released. However, if a study that was already reviewed by the WWC is included in a version 5.0 synthesis, it will be rereviewed using the standards described in the *WWC Procedures and Standards Handbook, Version 5.0*. Furthermore, all version 5.0 reviews will use the same protocol, [Study Review Protocol, Version 5.0](#), which means that a study would need to be reviewed only once under the *WWC Procedures and Standards Handbook, Version 5.0*.

2. When will reviews for grant competitions begin using the *WWC Procedures and Standards Handbook, Version 5.0*?

There is a gap between when the *Handbook* is released and when it is adopted for reviews conducted for U.S. Department of Education (ED) grant competitions. Reviews for the grant competitions in 2022 will be conducted using version 4.1 of the *WWC Procedures Handbook* and the *WWC Standards Handbook*. The WWC expects to continue to use version 4.1 standards to assess evidence cited for fiscal year 2023 grant competitions. The WWC expects to use version 5.0 standards for other reviews of studies, including systematic reviews, beginning in calendar year 2023.

3. How will the new standards affect the study rating for Education Innovation and Research (EIR) grants that are either already funded based on version 4.1 standards or are funded beginning in January 2023?

Grantees should consult their project officer for specific advice about their project.

4. Are any major updates planned to the Version 5.0 Online Study Review Guide?

The WWC is updating the Online Study Review Guide to align with the changes reflected in the *WWC Procedures and Standards Handbook, Version 5.0*. Users can expect updates to the Online Study Review Guide, including updated reviewer guidance, to be available by spring 2023.

5. Will certain protocols be assigned to version 5.0 or remain with version 4.1?

As of January 2021, the WWC uses the *Study Review Protocol* alongside the *WWC Handbook* to guide study reviews. A new version of this protocol, [Study Review Protocol, Version 5.0](#), has been released for public comment. This protocol will guide all study reviews under version 5.0.

The WWC will continue to develop topic area synthesis protocols to define the scope of new systematic reviews using the *WWC Procedures and Standards Handbook, Version 5.0*. Topic area synthesis protocols will 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.

6. When will reviewers have access to the version 5.0 training and certification activities?

Updated training videos for the *WWC Procedures and Standards Handbook, Version 5.0*, will be available in early 2023, along with updates to the WWC's study review guides. Once the training videos are released in early 2023, trainees can begin the process to become certified WWC reviewers by viewing all training video modules. Once trainees have completed all the training video modules, those who are interested in becoming certified WWC reviewers will be able to access and complete the certification exams.

7. Why do experienced reviewers need to certify again?

When the WWC moves to the next version of procedures and standards, certification requirements must be updated to align with the current version. Because of the volume and complexity of the updates to the procedures and standards in the *WWC Procedures and Standards Handbook, Version 5.0*, all individuals interested in obtaining certification as a WWC reviewer will need to complete the certification process. We anticipate that the training and certification materials will be available in early 2023.

8. Why is the role of topic area protocols reduced under version 5.0?

Under previous versions of WWC procedures, decisions that were delegated to topic area teams have either been pushed down to review teams or pushed up to the *Handbook* itself. Having most of the WWC's standards available in the *Handbook* increases review consistency, efficiency, and transparency. At the same time, some aspects of the standards depend on factors observed in individual studies, which implies that review teams are in the best position to make the necessary judgments. Furthermore, the WWC envisions that having review teams make some of these application decisions will result in more transparency about the reasoning behind those decisions.

9. Doesn't pushing decisions to review teams, rather than making them across the WWC, make different ratings more likely?

Under previous versions of the WWC procedures, multiple study reviews could result in multiple study ratings. Under the *WWC Procedures and Standards Handbook, Version 5.0*, the vision is that each study will have one review. The WWC plans on establishing processes to adjudicate disagreements if, for example, a later review team disagrees with judgments made by an earlier review team.

10. Why doesn't the WWC include supplemental findings in the effectiveness ratings?

An effectiveness rating signals whether an intervention did or did not impact a change in outcomes. The WWC relies on the main findings to determine effectiveness ratings because they have more credibility than supplemental findings in determining impact. They are usually based on a full study sample, capture impact closest to the end of the intervention, and, among other characteristics, are measured as a composite score instead of subscales. Supplemental findings can play a role in effectiveness ratings, however. For example, the WWC may pool supplemental findings to construct a main finding. For synthesis products, such as intervention reports, the topic area team may allow supplemental findings to contribute to the review. For example, findings for English learner subgroups would usually be considered supplemental findings. However, if the focus of a synthesis is on English learners, then the topic area protocol will indicate that findings for this subgroup are eligible for inclusion in synthesis products.

11. How is the effectiveness rating assigned to a study with more than one domain?

A study or intervention report receives the highest rating of evidence among its outcome domains. For example, if the WWC determined that a study demonstrated strong evidence in one outcome domain and uncertain effects in another outcome domain, then the study or intervention report will be listed as showing strong evidence in at least one outcome domain.

12. Can you provide more detail about the weighting of findings in synthesis products to allocate 51 percent to findings rated *Meets WWC Standards Without Reservations*?

Under version 5.0, if the WWC needs to synthesize across findings that meet WWC standards, then the WWC will attempt to attribute most of the meta-analytic weight to the findings rated *Meets WWC Standards Without Reservations*. This procedure will help the WWC ensure that the meta-analytic average is based primarily on effect sizes from findings that most credibly facilitate causal inference. The WWC will implement this reweighting procedure only when *all* the following conditions are met:

- The synthesis includes one or more findings rated *Meets WWC Standards With Reservations* and one or more findings rated *Meets WWC Standards Without Reservations*.
- Findings rated *Meets WWC Standards With Reservations* account for more than 50 percent of the default inverse-variance meta-analytic weight.
- The total sum of default inverse weights for findings rated *Meets WWC Standards Without Reservations* is 87.2 or greater, which aligns with the minimum level of precision for detecting an effect size of 0.30 standard deviations with 80 percent power.

If any one of these conditions are not met, then the WWC will synthesize the effect sizes from the corresponding findings using the default inverse-variance weights alone. However, the WWC will conduct a moderator analysis to statistically test the differences in effect sizes from the two WWC ratings. If the effects are statistically significantly different, then a cautionary note indicating the results of the moderator analysis will accompany the aggregate meta-analytic findings. Additional details about this procedure are included in appendix F in the *WWC Procedures and Standards Handbook, Version 5.0*.

13. Under the new definition of a study, how are site-specific impact estimates reviewed, and does it matter if there is a pooled impact estimate or not (for example, if an intervention was studied in two communities by the same research team and separate impact estimates were created [with or without a later pooled estimate])?

If site-specific impact estimates are included in separate manuscripts and not pooled elsewhere, the WWC will report these estimates as separate main findings. If multiple site-specific impact estimates are included in the same manuscript, then the WWC will report the pooled cross-site finding as the main finding. In any synthesis product, such as an intervention report, the WWC will report the pooled cross-site finding, whether it is included in separate manuscripts or a single manuscript.

14. Under what scenarios should a reviewer query to determine dependent samples?

It may be unclear whether two studies use dependent samples. To facilitate this determination, review teams should compare the following fields across the samples in question:

- Full reference. Each author and their Open Researcher and Contributor ID (ORCID), up to seven names.
- Intervention name.
- Year(s) in which students received the intervention.
- Year(s) in which outcome data were collected.
- Each location in which a study was performed. The WWC defines locations as a district or state because exact school locations are rarely available in manuscripts.
- Grade level of participants in the sample.
- Number of participants in the sample.

If the contents of some of these fields for a manuscript overlap with or are the same as those of another manuscript, then the review team may query the study authors to determine whether the studies have overlapping samples. If study authors do not respond to the query, then review team leadership will make a judgment based on the available information.

15. Can you clarify the role of the second reviewer and reconciler, and whether an author query (AQ) is necessary? Is there any scenario where a second reviewer would be assigned before sending an AQ?

The WWC does not mandate a process for contractors regarding the timing of the assignment of a second reviewer. If an AQ is needed to determine if a study will meet WWC standards, then it is generally advisable to assign the second reviewer after receiving a response to the AQ.

16. Under what scenarios should we query for contextual information about the study sample or intervention?

AQs attempt to gather missing information needed to determine a study's eligibility, the WWC research rating for each finding in a study eligible for WWC review, an acceptable effect-size estimate, or contextual information, such as sample demographics. The WWC will query for contextual information for studies rated *Meets WWC Standards Without Reservations* or *Meets WWC Standards With Reservations*. Appendix B in the *WWC Procedures and Standards Handbook, Version 5.0*, summarizes procedures for sending AQs.

Questions about WWC standards

17. Could you clarify how the independence definition will apply to reviews under the *WWC Procedures and Standards Handbook, Version 5.0*?

A new procedure under version 5.0 allows the effectiveness ratings for individual studies to be based on outcomes that are independent of intervention developers and study authors. The WWC will make the independence determination for measures in the outcome domains specified in the *Study Review Protocol* by applying the definition provided in version 5.0 of the *Handbook* ([pp. 28-29](#)), which states that a measure will be considered nonindependent if it was developed by study authors and is not in broad use or if it was developed by the intervention's developers. A nonindependence determination will not impact a study's eligibility for review or the study rating.

18. If the same developers developed the intervention and the measure, could the measure ever be independent?

Measures created by developers to accompany an intervention are nonindependent when used to estimate the impact of the same developer's interventions. However, if this measure is used to assess the impact of a separate intervention from different developers, it could be considered independent.

19. Do the rules about measure independence apply equally to developers and evaluators?

The rules are the same for developers and evaluators. For example, if evaluators developed a measure for an intervention they are evaluating, the measure will be nonindependent.

20. How does measure independence apply to situations when some, but not all, study authors were involved in the measure development?

If any study author was directly involved in the measure development, the WWC will consider the measure nonindependent.

21. The *Handbook* states that only some outcome domains will be assessed for independence and that these will be listed in the *Study Review Protocol*. Which outcome domains can authors expect to be reviewed for independence?

The [Study Review Protocol, Version 5.0](#), will include a complete list of all the outcome domains for which the WWC will use independent measures to assess effectiveness; at present, all of the outcome domains that will be assessed for independence are in math or literacy. The WWC chose to focus on math and literacy outcome domains because these domains have a relatively plentiful number of recognized, widely accepted, and independent measures available for use.

22. Measures developed by intervention developers can be created to account for correlated measurement error and reduce the likelihood of it occurring. How will the WWC assess these measures for independence?

The WWC acknowledges that measures can be designed to limit the potential bias of nonindependence of the outcome measures with intervention developers or study authors. However, the WWC does not have procedures to assess the psychometric properties of these measures. Therefore, when an outcome domain has a plentiful number of independent measures, the WWC will base the effectiveness ratings in individual study reviews on measures independent of intervention developers and study authors.

23. How will the WWC review measures for independence?

The WWC will review measures according to the definition of independence written in the [Study Review Protocol, Version 5.0](#), and the WWC Procedures and Standards Handbook, Version 5.0: “a measure will be considered nonindependent if either it was developed by study authors and is not in broader use, or if it was developed by the intervention’s developers.”

When review teams encounter measures that meet the definition of independence as stated in the *WWC Procedures and Standards Handbook, Version 5.0*, but are not currently included in the list of known independent measures in the *Study Review Protocol, Version 5.0*, they will be able to consult the WWC for an independence determination by reaching out to the WWC Help Desk at <https://ies.ed.gov/ncee/wwc/help>. More information on this determination process will be available in the *Study Review Protocol, Version 5.0*.

24. Would researcher-made assessments for high school math (e.g., advanced algebra) be eligible for review as there are no widely available norm-referenced assessments?

Researcher-made assessments, which are referred to as “nonindependent measures” in the *WWC Procedures and Standards Handbook, Version 5.0*, are eligible for review.

The fact that a measure is nonindependent has potential implications for whether findings associated with that measure are designated as main or supplemental. The [Study Review Protocol, Version 5.0](#), will identify outcome domains for which the WWC will evaluate measure independence.

25. If a study modified an independent measure, will it still be considered independent? Will version 5.0 of the *Handbook* provide guidance on when and how to determine whether such measures are still considered independent?

If an independent measure is modified for a specific intervention, the WWC will consider the modified measure nonindependent for that intervention.

26. Why is only face validity, and not other types of validity, considered in the WWC outcome measure standards?

Other types of validity, including content, predictive, concurrent, and construct validity, are important considerations for measure developers when designing and validating an outcome measure in education. Because the WWC primarily reviews studies that measure the effectiveness of policies, practices, and programs and not the development of outcome measures, the WWC does not assess outcome measures for all types of validity.

27. What are the conditions under which high-attrition randomized controlled trials (RCTs) will not require a baseline equivalence assessment? Why did the WWC make this change?

The WWC will no longer assess baseline equivalence in high-attrition RCTs or studies that use regression discontinuity design (RDD) when attrition bias is assessed using the optimistic boundary. The WWC made this change in part due to data demonstrating that most RCTs that are labeled “high attrition” are RCTs in which the WWC is unable to assess the level of attrition due to incomplete information available at the time of review. If the WWC were able to assess the level of attrition in these RCTs, it is likely that they would represent a mix of low- and high-attrition RCTs. In addition, the WWC allows this flexibility because although 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 reviews of RCTs and studies that use RDD into closer alignment than in previous versions of the standards.

28. Can you clarify the closer alignment with Every Student Succeeds Act (ESSA) evidence tiers, such as Tier 3 evidence? Do the standards discuss what is required to meet that tier?

The WWC assigns an evidence tier designation according to ED’s evidence definitions. See the video titled [Using the WWC to Identify ESSA Evidence Ratings](#) to learn more about how the WWC aligns evidence tiers to studies.

29. Can you clarify whether “negative effects” and “uncertain effects” will receive effectiveness ratings?

The WWC uses effectiveness ratings to communicate whether an intervention did or did not affect a change in outcomes, and “negative effects” and “uncertain effects” are types of effectiveness ratings. When the main study findings in an outcome domain are not statistically significant or unknown, or if the study has at least one supplemental finding that meets standards and no main findings, the outcome domain will receive an effectiveness rating of “uncertain effects.” If the main study findings in an outcome domain are statistically significant but negative, the outcome domain will receive an effectiveness rating of negative effects. [Table 23](#) and [Table 24](#) in the *WWC Procedures and Standards Handbook, Version 5.0*, provides additional information on the criteria used to assign effectiveness ratings.

30. The *Handbook* states, “Student behavior based on classroom observations also will be considered as arising from cluster-level assignment (with students as the unit of measurement), even if the study authors recorded their observations at the classroom level.” I take this to mean that measures like the ones we’ve seen in single-case design studies that are measured for classrooms and are not strictly aggregates of student-level assessments would be considered as measured at the student level. I see some issues with this approach, but the main thing is to understand where the line is between a classroom- and student-level measure.

The *WWC Procedures and Standards Handbook, Version 5.0*, makes the distinction between student and teacher outcomes regarding measures based on classroom observations. The current *Study Review Protocol* that applies to version 4.1 procedures and standards further elaborates on this distinction by defining specific outcome domains, such as “Student Behavior” and “Teacher Practice.” Although measures in both domains could come from classroom-level observations, they differ in terms of whether students or teachers are the focus of the observations. Hence, in a teacher-level RCT, analyses of teacher practice outcomes would be considered individual-level assignment, whereas analyses of student behavior outcomes would be considered cluster-level assignment, even if both sets of outcomes were based on classroom observations.

General questions

31. How do the changes to the WWC standards on measure independence relate to Mark Schneider’s remarks?

By prioritizing broadly accepted independent measures when they are available, the changes to the WWC standards on measure independence align with the director's remarks.

32. We started a comprehensive meta-analysis with the previous WWC version. We have completed screening and will begin the design evaluation phase. Would you suggest making changes?

Not necessarily; however, study authors should consult the finalized *WWC Procedures and Standards Handbook, Version 5.0*, that was published in August 2022.

33. What is the timeline for making extracted data available from reviews for interested researchers (for example, extracted data from single-case design studies)?

After a review team finalizes a study review and publishes it on the public WWC website, the WWC, at the same time, publishes relevant summary data from that study review to the [Data From Study Reviews Database](#), which can be downloaded for further analysis.

Related resources

In addition to the webinar and this Q&A document, the following resources may be useful:

- The *WWC Procedures and Standards Handbook, Version 5.0*: <https://ies.ed.gov/ncee/wwc/Handbooks>
- Summary of changes to the *WWC Procedures and Standards Handbook, Version 5.0*: https://ies.ed.gov/ncee/wwc/Docs/referenceresources/Final_HandbookSummary-v5-0-508.pdf
- The *Study Review Protocol, Version 5.0*: <https://ies.ed.gov/ncee/wwc/Document/1297>
- WWC Review Protocols: <https://ies.ed.gov/ncee/wwc/Handbooks#protocol>
- WWC Help Desk: <https://ies.ed.gov/ncee/wwc/Help>
- Data From Study Reviews Database: <https://ies.ed.gov/ncee/wwc/studyfindings>
- Sign up to receive Institute of Education Sciences news briefs: <https://ies.ed.gov/newsflash/>.