

Questions and Answers About Changes in WWC Version 4.1 Standards and Procedures

The purpose of this document is to provide answers to the questions submitted before and during the What Works Clearinghouse (WWC) technical assistance webinar, "What's New in Version 4.1 of the WWC Standards and Procedures Handbooks?" The webinar was hosted on January 22, 2020.

This document is meant to serve as a companion to the webinar slide deck and webinar recording, which can be found here-on-the-wwc-website. 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 STANDARDS AND PROCEDURES

1. If the overall meta-analytic average is statistically significant, but derives from two studies, one meeting standards without reservations and one with reservations, what is the intervention rating?

The intervention rating depends on the percentage of the meta-analytic average that is attributable to the study rated *Meets WWC Standards Without Reservations*. The *WWC Procedures Handbook*, *Version 4.1*, articulates that for an intervention to have a rating of "positive effects," more than 50% of the fixed-effects meta-analytic weight must come from studies that are rated *Meets WWC Standards Without Reservations*. This means that if there are only two studies in a meta-analysis, and if the study rated *Meets WWC Standards Without Reservations* has the larger sample size, the intervention will have a rating of positive effects.

2. Within the fixed-effects method, what approach is used when effect sizes from more than one contrast are available for inclusion in the meta-analytic weight (e.g., a multi-arm intervention study)? Should one contrast be chosen for inclusion, or is there another methodological option to allow inclusion of multiple dependent effect sizes from one study?

Currently, the WWC does not have specific guidelines for studies that use multiple comparison groups. Review teams have approached these studies by (a) including all comparisons they consider relevant, (b) calculating separate effect sizes for each comparison, and (c) averaging these findings together in a manner similar to multiple outcomes in a domain. The lead methodologist should use discretion to decide the best approach for the team on a study-by-study basis.



3. Is the effect size generated from the design-comparable effect size (D-CES) on the same scale as an effect size generated from a group design?

The effect size estimate generated from the D-CES is in standard deviation units, which is the same scale as the standardized mean difference effect size generated from a group design study.

4. Is the idea of "substantive effects" removed in 4.1?

The handbooks no longer include the concept of "substantively important" for effects of 0.25 standard deviations or larger that are not statistically significant.

5. Why was the substantively important designation removed? And do you know if other guidelines that other agencies may use have done this as well?

Prior to Version 4.1 of the WWC Standards and Procedures, an effect size of 0.25 standard deviations or above was determined to be "substantively important." The removal of the "substantively important" designation was motivated by the WWC's decision to revise procedures for assigning effectiveness ratings to interventions. The new system relies on a fixed effects meta-analysis of effects from individual studies and the assessment of the sign and statistical significance of that average effect, rather than on a vote-counting approach relying on individual studies suggesting positive effects. The new emphasis on pooling effect size estimates across studies increases the power of statistical hypothesis tests and does not incorporate judgments about the magnitude of the observed effects into judgments about intervention effectiveness. The U. S. Department of Education's evidence definitions adopted in 2017 also emphasize an effect's sign and statistical significance, rather than its magnitude, as a necessary criterion for strong, moderate, or promising evidence of effectiveness. For these reasons, the "substantively important" designation was no longer necessary or helpful for the WWC to use in its reviews and syntheses of evidence.

Additional information is available in Responses to Comments From the Public on Updated Version 4.1 of the *WWC Procedures Handbook* and *WWC Standards Handbook*: https://ies.ed.gov/ncee/wwc/Docs/referenceresources/SumResponsePublicComments-v4-1-508.pdf.

6. What does 4.1 say about what Cook and Campbell called recurrent institutional cycle design and people like Stockard referred to as using cohort control groups? Prior WWC Standards were fairly broad in excluding any cohort comparison as valid evidence of effects, presumably including the sequential cross-cohort comparisons where entering youth scores were compared to graduating youth scores as a betweengroup difference test.

When the time period differs for the groups, time is a confounding factor. A design in which groups are defined by cohort is often labeled a successive-cohort design or cohort design. For example, an intervention group consists of a cohort of third graders in Year Y, and the comparison group consists of the previous cohort of third graders in Year Y–1. Usually, both cohorts are observed in one school or the same set of schools. In this cohort



design, the intervention and comparison conditions are completely aligned with different time periods, and the estimated impact is confounded with any changes that occur between those time periods. These changes—such as new district policies, new personnel, or new state tests—could plausibly affect outcomes. Because many of the changes that occur over time are likely to be unobserved or not reported, the WWC cannot assess how problematic the potential changes are in individual studies.

QUESTIONS ABOUT THE ONLINE STUDY REVIEW GUIDE (OSRG)

7. Will the OSRG reporting reflect the use of the fixed effects approach for synthesizing effect sizes in the same outcome domain? And If so, when?

The OSRG used by WWC certified reviewers will reflect all Version 4.1 updates, including the calculation related to synthesizing effect sizes in the same outcome domain. This OSRG will become available in the summer or fall of 2020. The public OSRG also will reflect all Version 4.1 updates and will be available by the fall or winter of 2020.

8. Will the reviews indicate which criteria were used to review?

When conducting reviews of studies, the reviewer is asked to indicate which versions of the Standards and Procedures were applied as well as the study review protocol. This information is available to the public through the <u>Review of Individual Studies</u>.

9. Are WWC reviewers to use the Pustejovsky software to calculate D-CES? Or are there D-CES calculation procedures in the OSRG?

Currently, a D-CES can be calculated for multiple baselines across cases, multiple probes across cases, and treatment/reversal/withdrawal designs with three or more cases. Decisions related to the calculation of D-CES will be guided by the review team leadership. For this reason, the calculation will not be built into the OSRG. For illustrative purposes, the WWC training team uses the scdhlm web-based calculator (Pustejovsky, 2016), which is a user-friendly tool located online at https://jepusto.shinyapps.io/scdhlm.
The WWC anticipates calculating a D-CES for other types of single-case designs (SCDs) as those become available in the methodological literature.

¹ Pustejovsky, J. E. (2016). *Scdhlm: A web-based calculator for between-case standardized mean differences* (Version 0.3.1) [Web application]. Retrieved from https://jepusto.shinyapps.io/scdhlm



QUESTIONS ABOUT CERTIFICATION

10. I am a certified WWC reviewer under Version 4.0 of the handbooks. Do I have to complete any new activities to be certified to conduct WWC reviews under the Version 4.1?

Reviewers certified under Version 4.0 of the WWC Standards and Procedures will need to be recertified under Version 4.1. Reviewers who wish to recertify to Version 4.1 can learn more about the recertification process at

https://ies.ed.gov/ncee/wwc/onlinetraining/recertification. Please subscribe to the IES newsletter at https://ies.ed.gov/newsflash/ to be alerted when the video is available.

11. I started certification for group design standards under 4.0 but haven't completed yet. Do I need to start over?

You have two options. The online group design training modules and certification process will transition to Version 4.1 this spring. Your first option is to complete the certification process before then, in which case you will be certified in Version 4.0 of Standards and Procedures. Your next step will be to recertify under Version 4.1. You can learn more about the Version 4.1 recertification process at https://ies.ed.gov/ncee/wwc/onlinetraining/recertification.

Your second option is to wait until the *updated* group design modules are available in spring 2020. Then you can watch the updated modules. When you are done, you can take the updated Version 4.1 certification examination to be certified under Version 4.1.

12. I am certified in group design standards under the Version 3.0. What should I do to certify under 4.1?

If you have not recertified from Version 3.0 to Version 4.0 by January 31, 2020, then you will be required to complete the entire Version 4.0 certification process outlined at https://ies.ed.gov/ncee/wwc/onlinetraining. After you complete Version 4.0 certification, you will be able to recertify to Version 4.1. You can learn more about the Version 4.1 recertification process at https://ies.ed.gov/ncee/wwc/onlinetraining/recertification.

13. Will the video for the 4.1 certification be different from the current webinar recording?

The Version 4.1 recertification video differs from the webinar. The video is available at https://ies.ed.gov/ncee/wwc/onlinetraining/recertification.

14. What is the recertification process for SCD?

The WWC will not have a recertification process for SCD under Version 4.1. Instead, reviewers who wish to become certified in Version 4.1 of the SCD standards must attend an in-person training, pass a certification examination, and complete a study review.



15. Will studies previously reviewed under Version 4.0 of the handbooks be re-reviewed under Version 4.1?

Studies reviewed under Version 2.0 or earlier will be re-reviewed if they are included in a new WWC product or submitted for evidence as part of a grant competition. Studies reviewed under later versions, including Version 4.0, will not be re-reviewed. The WWC expects to review studies using Version 4.1 later in 2020.

General Resources

In addition to the webinar and this Questions and Answers document, the following resources provide guidance about the WWC's Standards and Procedures for reviewing evidence from regression discontinuity design studies. These include resources shared in the chat box during the webinar.

- WWC Standards Handbook, Version 4.1: https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf.
- WWC Procedures Handbook, Version 4.1: https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Procedures-Handbook-v4-1-508.pdf.
- Summary of Changes to WWC Procedures and Standards in the Version 4.1 Handbooks: https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWCHandbookSummary-v4-1-508.pdf.
- Responses to Comments From the Public on Updated Version 4.1 of the WWC Procedures
 Handbook and WWC Standards Handbook:
 https://ies.ed.gov/ncee/wwc/Docs/referenceresources/SumResponsePublicComments-v4-1-508.pdf.
- WWC Reviews of Individual Studies: https://ies.ed.gov/ncee/wwc/ReviewedStudies#/OnlyStudiesWithPositiveEffects:false,Set

 Number: 1
- WWC Group Design Standards Online Training: https://ies.ed.gov/ncee/wwc/onlinetraining.
- WWC Group Design Recertification: https://ies.ed.gov/ncee/wwc/onlinetraining/recertification.
- Sign up to receive Institute of Education Sciences news briefs: https://ies.ed.gov/newsflash/.