

Questions and Answers About Confounding in Group Design Studies

The purpose of this document is to provide answers to the questions submitted before and during the WWC technical assistance webinar hosted on April 18, 2019 about confounding in group design studies.

This document is meant as a companion to the slide deck used during the webinar, which can be found on the [webinar archive page](#). Whenever possible, we identified the slide(s) to which a question pertains. We combined similar questions and rephrased some others for clarity, preserving the meaning of the original question. If additional questions arise, please contact the WWC Help Desk at <https://ies.ed.gov/ncee/wwc/help>.

The end of this document provides links to resources mentioned throughout this document such as the WWC Standards Handbook 4.0, WWC Procedures Handbook 4.0, WWC Standards Brief, WWC Certification Module on Confounds, and WWC database of reviewed studies.

1. Does $n = 1$ confound extend to the district level? (Question applies to slides 6 and 16.)

The WWC considers a study unit to be an $n = 1$ confound if it (1) is observed, (2) aligns with only one study condition, and (3) is not part of the intervention (see WWC Procedures and Standards Handbooks, version 4.0, page 80 and slide 6). If a district in a study meets all three of these criteria, then the district will be considered an $n = 1$ confound. For example, on slide 16 in the presentation slide deck we present two scenarios in which a district constitutes an $n = 1$ confound. In one scenario, the intervention condition includes only one district (district A); in the other scenario, the comparison condition includes only one district (district C). One district in each scenario is an $n = 1$ confound because the district is observed, aligns with only one study condition, and is not part of the intervention.

2. What if multiple districts in the comparison and treatment conditions are very similar? For example, they are in the same urban area? Does the similarity of the district demographics or other factors have any effects on counting them as one or multiple clusters? (Question applies to slide 16.)

The WWC will not consider a study to have an $n = 1$ confound at the district level if multiple districts are represented in the intervention and comparison conditions. We illustrate this scenario on the right side of slide 16 in the presentation slide deck. Because the intervention condition in this scenario includes districts A, B, C, and D, and the comparison condition includes districts E and F, the study does not have an $n = 1$ confound at the district level. No single district is perfectly aligned with one study condition. If the districts in each condition are similar (e.g., in the same urban area as the author of the question proposed), then the WWC will

still consider the study to be free of an $n = 1$ confound at the district level because there are multiple districts in each condition.

3. Is the implementer necessarily a confound (if there is one implementer)? What if the same implementer works with both the intervention and comparison groups?

When evaluating a study for an $n = 1$ or characteristics confound, it is important to adhere to the WWC's definition of a confound: It is (1) observed, (2) aligns with only one study condition, and (3) not part of the intervention. An implementer constitutes an $n = 1$ confound if the implementer works with students in one condition but not the other (e.g., as in the study we describe on slide 22). If the implementer works with students in both conditions, then the implementer will not be considered an $n = 1$ confound.

4. Can demonstration of baseline equivalence (across plausibly relevant factors) ever assuage concerns of a time confound?

When the intervention and comparison groups are completely aligned with different time periods, the WWC will consider time to be a confound. If the groups are similar at baseline, the WWC will still consider time to be a confound because the impact of an intervention may be confounded with any changes that occurred between the time periods (see WWC Standards Handbook 4.0, pp. 81–82). On slides 41 and 42, we use a study example that includes an intervention group of seventh graders in 2006–07 and a comparison group of seventh graders from the same school in 2003–04. Even if the seventh graders from both groups were equivalent at baseline on characteristics specified in the review protocol, the study still would have had a time confound because the impact of the intervention could have been confounded by changes that occurred between 2003–04 and 2006–07.

5. Regarding an example in the combined interventions section: Why wouldn't just the intervention schools using the Connected Mathematics Project be eligible for review? Assuming the authors disaggregated the intervention schools by treatment condition. (Question applies to slides 53 and 54.)

In the example we use on slides 53 and 54, the authors did not disaggregate the intervention schools by type of math program (Math in Context and Connected Mathematics Project), but instead reported the results for both math programs in aggregate. Because the Connected Mathematics Project intervention was the only intervention eligible for review under this WWC review protocol, in this example, the combined intervention was considered a confound and the study was found *Ineligible for WWC Review*. Should the same study be reviewed under a different WWC protocol where the combined intervention was eligible for review, this combined intervention would not be considered a confound.

6. Can you briefly discuss the distinction between (a) being *Ineligible for WWC Review* and (b) being rated as *Does Not Meet WWC Group Design Standards*?

The first step in any WWC review is an initial screening of a study for eligibility against parameters specified in the review protocol. Studies may be found *Ineligible for WWC Review* for several reasons, such as not using an eligible design, not using a sample aligned with the protocol, or not being published within the timeframe specified in the protocol. Studies may also be found ineligible if the intervention examined in the study is not aligned with the focus of the protocol. Such a situation may arise if a study examines the effectiveness of a combined intervention—a special case of confounding that we discussed during the webinar. A study that examines a combined intervention may be found *Ineligible for WWC Review* if the focus of the review protocol is on one intervention examined in the study, not both.

Once a study is determined *Ineligible for WWC Review*, the study review stops at this stage, and the study does not receive a WWC rating. If a study is eligible for review, the review proceeds, and a WWC-certified reviewer assesses the study against WWC group design standards. At this stage the study will receive one of three ratings: *Meets WWC Group Design Standards Without Reservations*, *Meets WWC Group Design Standards With Reservations*, or *Does Not Meet WWC Group Design Standards*. Studies that contain an $n = 1$ confound or the characteristics confound receive the rating of *Does Not Meet WWC Group Design Standards* (assuming they pass the initial screening and are found eligible for review). The WWC Procedures Handbook 4.0 describes the review process in detail (e.g., see pp. 7–8 for determining study eligibility).

7. The blended learning example caused some confusion. Is it the whole school intervention (which it sounded like) or is it just the “site”? (Question applies to slides 28 and 29.)

The study described on slides 28 and 29 was conducted in the FirstLane charter organization’s four elementary schools. One school implemented the blended learning model in grades K–8 for one academic year. This school used blended learning to supplement the Response to Intervention program and instruction for students with special needs (we updated slide 28 to indicate that select students were receiving the intervention). Three other schools served as the comparison condition. The purpose of the study was to examine the effectiveness of the blended learning intervention on mathematics and English language arts learning. The intervention was delivered in only one school. Because the school and the intervention condition are perfectly aligned, the school constitutes an $n = 1$ confound. Although the blended learning approach was used in all grades in the intervention school, it is not the same as the school itself being an intervention. When the school and the intervention are the same, attending the school equates to receiving the intervention. Instead, in our example the school was a site for delivering the blended learning model for one year to a select group of students. We describe the idea of the school being the same as the intervention more fully in response to the question 8.

8. Can you explain why when the school is the intervention it is not an $n = 1$ confound, even if there is only one school? The distinction of the whole school being the intervention is not clear. It still seems the same as the school being the site for the intervention. (Question applies to slide 17.)

If the school and the intervention are the same, the WWC will not consider one school to be an $n = 1$ confound, even when it is the only school contributing students to the intervention condition. If the school is the intervention, then being at the school, or attending the school, equates to receiving the intervention. In the career academy example used in the webinar (slide 17), the career academy is one whole school and attending the schools is the same as receiving the intervention. An example provided in the WWC Standards Handbook 4.0 is of a school with a unique organization and governance compared with other traditional schools. Attending the school with a unique governance is the same as receiving the intervention, which is why in this example the one school is not an $n = 1$ confound (see p. 81). A single school becomes a confound when an intervention is implemented within the school, but the school and the intervention are not the same (i.e., the school is a site for an intervention). For example, if a study examines the effectiveness of a new curriculum on math achievement of seventh graders, with all students in the intervention group from the same school, then a single school in the intervention group will be an $n = 1$ confound. In this example, attending the school is not the same as receiving the intervention. The school is a site for an intervention administered to seventh graders. The effect of the intervention (the new math curriculum) cannot be separated from the effect of one school in the intervention condition.

The examples used during the webinar and in the above paragraph do not take a review protocol into account. However, it is important to note that for any WWC study review, a review protocol will dictate whether a single school constitutes an $n = 1$ confound. For example, if the review protocol stipulates that an eligible intervention must be replicable, then a single school, even if the school and the intervention are the same, will be a confounding factor because a single school cannot be replicated.

9. I'm having trouble understanding the example of all the seventh graders in the comparison group were ranked second highest in math. How is it that this is not considered a characteristic confound? (Question applies to slide 41).

On slide 41 we describe a study that includes an intervention group of students from 2006–07 and a comparison group of students from 2003–04. The author notes that students in the intervention group were the “second highest groups of math students” in the school based on state standardized test scores. The level of math achievement of students in the intervention group may be a characteristics confound if students in the comparison group do not have the same level of math achievement (i.e., there is no overlap on math achievement across the two groups). According to the reviewer, the study does not provide information about math achievement of students in the comparison group. Had the level of math achievement in the intervention and comparison groups been the only concern about the study’s design, an author query would have been required to clarify group characteristics. However, another concern about

the study's design is a time confound: Students in the intervention and comparison groups were observed in different school years. The presence of time confound is unambiguous. Therefore, the only rating the study can receive is *Does Not Meet WWC Group Design Standards*. The presence or absence of another confound based on math achievement would not change the rating. If receiving additional information would not change the rating, an author query would not be sent.

10. We received several questions about specific research-based practices meeting the WWC evidence requirements, for example Marzano's practices.

Only interventions that have been reviewed by the WWC will have a WWC rating of effectiveness. People interested in a specific study or intervention should search the WWC database of reviewed studies. If a study or intervention does not appear in the database, it means that WWC has not reviewed them. It does not mean that the study or intervention did not meet the WWC evidence standards.

11. We received several questions about how to design and implement studies that are free of confounds.

The best approach is to think through the study design carefully during the planning stage. At the planning stage, researchers should ensure that the study has enough units in all conditions, and that the characteristics of groups and participants are shared between conditions. Even with the most careful planning, a study can have unexpected developments that can result in a confound. For example, attrition can lead to only one study unit (e.g., one intervention school) available for the analysis of outcomes, which would constitute an $n = 1$ confound. Remember, however, that the WWC rating is not intended as the rating of the usefulness or value of research but rather is a rating of internal validity. Finally, keep in mind that if you plan a study that combines interventions, depending on the WWC protocol being used to review a study, the study may or may not be eligible for review. Research developers interested in obtaining the highest possible WWC ratings should, at the research planning stage, identify the protocol under which a study might be expected to be reviewed.

12. Does confounding apply to group interventions?

A confound can occur in any intervention, including group interventions. For example, a study that assigns groups to conditions can have an $n = 1$ confound if only one group is included in either condition. The "group" can be a school, a classroom, or any other collective. A confound can also occur at any level. For example, a study may include multiple schools in both conditions, but only one teacher delivers the intervention. This study does not have an $n = 1$ confound at the school level, but it has an $n = 1$ confound at the instructor level.

13. Are there templates and exemplars for district researchers to use?

The WWC does not have one solution or a template for every study. However, in the webinar we discussed a filter for deciding whether a study contains a confound. The filter includes three

criteria for determining an $n = 1$ or characteristics confound: Is it observed, is it aligned with only one condition, and is it part of the intervention (e.g., slide 13). Researchers can consider assessing study units and group/participant characteristics through the filter to decide whether they may be a confound. If you plan a study that combines interventions, then as discussed in the answer to question 11, identify the protocol under which your study will be reviewed. If only one intervention is the focus of the protocol, your best protection against a confound in that case is to deliver the second intervention to the comparison group. The best time to go through these steps is in the study planning phase.

14. How does the WWC definition of confounds compare with the definition of confounds in other disciplines?

Different disciplines have different unique characteristics of confounding factors. However, the unifying idea of a confound is that it is a design feature that prevents the researcher from accurately estimating the intervention effect.

General Resources

In addition to the webinar and Questions and Answers document, the following resources provide additional guidance about confounding in group design studies and other relevant topics.

- WWC Group Design Standards Handbook 4.0:
https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf
- WWC Procedures Handbook 4.0:
https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_handbook_v4.pdf
- WWC Standards Brief on confounding factors:
https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_brief_confounds_101117.pdf
- WWC Reporting Guide for Study Authors: <https://ies.ed.gov/ncee/wwc/ReportingGuide>
- WWC Certification Module on Confounds (including transcript):
https://ies.ed.gov/ncee/wwc/Docs/OnlineTraining/wwc_training_m4.pdf
https://ies.ed.gov/ncee/wwc/Docs/OnlineTraining/wwc_training_m4_transcript.pdf
- WWC database of reviewed studies: <https://ies.ed.gov/ncee/wwc>.