

Review Protocol FOR FRACTIONS PRACTICE GUIDE
(This guide is subject to [WWC version 2.0 standards](#))

The Fractions practice guide will provide evidence-based guidance on the teaching and learning of fractions among students in the elementary and middle grades. The guide will offer concrete recommendations for practices that improve students' conceptual understanding of fractions, operational fluency with fractions, and problem-solving skills with fractions.

SCREENING CRITERIA

Time Frame

Relevant to: Screening Criteria

Protocol: **1989–present** (specific highly-related studies from 1983–1988 may also be reviewed upon request by the panel)

Study Design

Relevant to: Screening Criteria

Protocol: **RCTs, QEDs, RDDs, and SCDs are eligible for review. Correlational and Qualitative studies will be held for future consideration.**

Regression Discontinuity Designs (RDDs) and Single-Case Designs (SCDs). RDDs and SCDs are eligible for review. Until WWC standards for RDD and SCD studies are disseminated, please note that the study is an RDD and the evidence coordinator will have it evaluated by an expert in RDDs or SCDs.

Eligible interventions and comparisons:

- Intervention groups may receive “bundled” interventions (i.e., the intervention may be multi-faceted)
- Studies that compare multiple interventions are eligible for review (i.e., the comparison group does not have to be “business as usual”)
- Studies comparing multiple levels of treatment are eligible (for example, Intervention A might be compared with Intervention A+B)

Completing SRGs with more than two groups. Studies may include three or more eligible groups (for example, Intervention A, Intervention B, and a comparison group). If this is the case, reviewers will evaluate whether the study meets standards for each contrast (i.e., pair of groups). If any contrast meets standards, the study will be rated as meeting standards (or meeting with reservations). The SRG and Study Tracker Notes should clearly indicate any contrasts that do not meet standards, or that have a different rating from the main rating indicated in the study tracker. When there are more than two groups, Reviewers should complete separate calculations on differential attrition for each contrast and separate effect size calculations for each contrast (Note: multiple comparison adjustments should be made across groups. See “Multiple Comparisons” below).

Correlational and Qualitative Research. Correlational and qualitative studies cannot be used to support a moderate or strong level of evidence for a recommendation, but may be used for recommendations with low levels of evidence that have few studies or no causal studies. Studies with these design may be flagged for later summarization, but will not be assigned as part of the formal review process.

Brief Description of Intervention Focus

Relevant to: Screening Criteria

Protocol:

The recommendations in the practice guide will focus on instructional strategies to develop students' proficiency in fractions, including their (1) conceptual understanding of fractions, (2) operational fluency with fractions, and (3) problem-solving skills with fractions. This includes the following concepts and skills (as defined by the National Math Advisory Panel):

- identifying and representing fractions and decimals;
- comparing, adding, and subtracting fractions and decimals;
- multiplying and dividing fractions and decimals; and
- conducting operations with positive and negative fractions, and solving problems involving percent, ratio, and rate.

Since fractions are part of the broader concept of rational numbers, the guide will also address related concepts such as ratios, decimals, proportions, and percentages.

Fluency with whole numbers, a necessary precursor to working with fractions, will not be addressed by the practice guide. The guide will not address the teaching of fractions to students with learning disabilities or to secondary students who require remediation.

Age and/or Grade Range, and Eligible and Ineligible Subgroups

Relevant to: Screening Criteria

Protocol: **Grades Kindergarten through 8th grade students**

- Studies of students in grades Kindergarten through 8th grade (ages 5 through 14), or in any subset of these grades are included.
- Studies that contain students in other grades are not included unless (1) study results disaggregate the results of students in eligible grades, or (2) students in eligible grades represent the majority of the aggregated mixed-age sample.

Focus on English-speaking students without identified disabilities. The panel will not focus on, nor will it issue recommendations centered on, comprehension instruction and practices that target English language learners or students with learning disabilities. Studies that exclusively target students with learning disabilities will not be included, but studies that target students with learning disabilities as well as students without learning disabilities will be included.

Location and Language

Relevant to: Screening Criteria

Protocol: The scope of the guide includes studies of fraction interventions that were administered in **English or in other languages**. Studies do not have to be set in the United States.

Outcomes and Domains for Multiple Comparisons

Relevant to: Screening Criteria AND Full Review

Protocol:

Eligible outcomes relate to children’s understanding and learning of fractions, as well as their math achievement more generally. These outcomes should be organized into one of the following categories and are treated as separate domains for multiple comparison adjustments:

Fraction knowledge—children’s conceptual understanding of fractions, decimals, and ratios; this includes fraction notation, part-whole relationships, knowledge of fraction magnitudes, and fraction representations.

Computation involving fractions, decimals, ratios, rates, or proportion; this includes addition, subtraction, division, multiplication, and other operations with fractions and related concepts.

Word problems involving fractions, decimals, ratios, rates, or proportion (if computation is embedded within a word problem, categorize as “word problems”).

Math achievement tests that measure students’ overall math knowledge and skills (i.e., not necessarily focused on fraction learning).

Studies with Multiple Groups: Studies may include more than two groups. If that is the case, when adjusting for multiple comparisons, reviewers should count the total number of outcome/pair combinations for a given intervention. For example, consider a study that has five outcomes in a domain, and three groups (Intervention 1, Intervention 2, Comparison), where all groups have data on all five outcomes. On a spreadsheet calculating the effect of Intervention 1, the total number of groups for a multiple comparison adjustment will be 10: I1 vs I2—five outcomes **plus** I1 vs Comparison—five outcomes. (If your study contains multiple age groups, some of which are outside the protocol, please discuss the correct adjustment procedure with your reconciler or the Evidence Coordinator.)

FULL REVIEW

Attrition Standards for RCTs

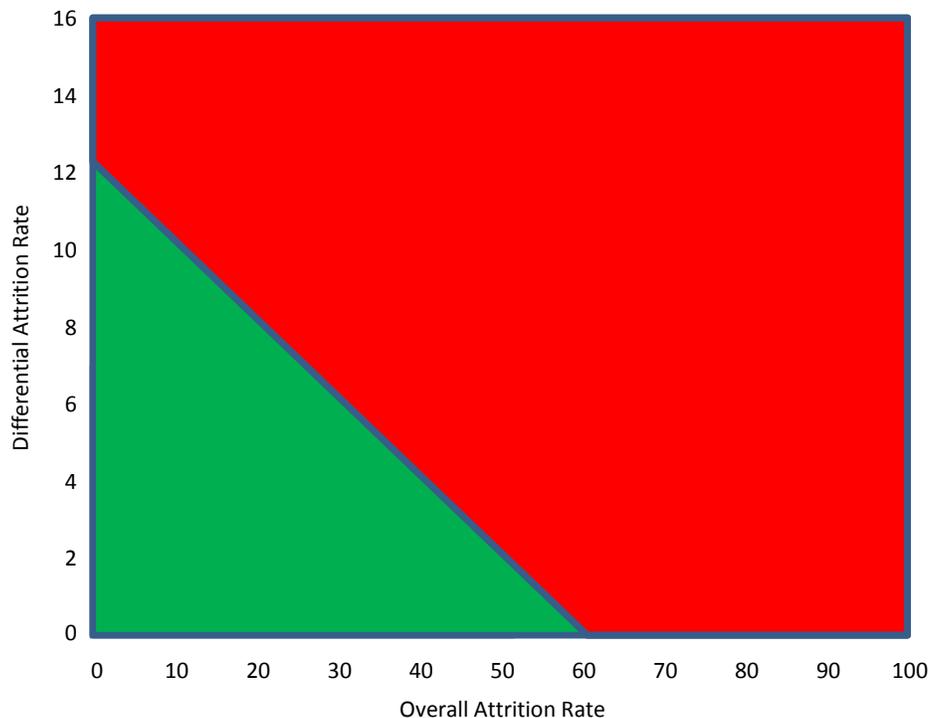
Relevant to: Full Review

Protocol:

As described in the *WWC Procedures and Standards Handbook (version 2.0)*, the WWC is concerned about overall and differential attrition from the intervention and comparison groups for RCTs, as both contribute to the potential bias of the estimated effect of an intervention. The attrition bias model developed by the WWC will be used in determining whether a study meets WWC evidence standards (see Appendix A of the *Handbook*).

When the combination of overall and differential attrition rates causes an RCT study to fall in the green area on the diagram shown below, the attrition will be considered “low” and the level of bias acceptable. The Fractions guide is using the “Liberal” attrition rule in the SRG. This reflects the assumption that most attrition in studies of fractions is due to factors that are not strongly related to intervention status, such as parent mobility and absences on the days that assessments are conducted. However, for RCTs with combinations of overall and differential attrition rates in the red area, the attrition will be considered “high” and potentially have high levels of bias, and therefore must demonstrate equivalence.

Note the WWC currently assesses attrition at both the cluster and sub-cluster levels.



Baseline Characteristics for QED Equivalence (and RCTs with high attrition)

Relevant to: Full Review

Protocol:

If the study design is an RCT with high levels of attrition or a QED, the study must demonstrate baseline equivalence of the intervention and comparison groups for the analytic sample. The onus for demonstrating equivalence in these studies rests with the authors. Sufficient reporting of pre-intervention data should be included in the study report to allow the review team to draw conclusions about the equivalence of the intervention and comparison groups.

QEDs and RCTs with high levels of attrition must demonstrate baseline equivalence (0.25 standard deviation or less) of the analysis sample on pretest measures of the outcomes in order to meet standards with reservations. Failure to demonstrate baseline equivalence on a measure of fraction knowledge or skills, general math achievement, or another appropriate measure will cause the study to fail to meet standards. If differences are greater than 0.05 standard deviations and less than or equal to one-quarter of the pooled standard deviation in the sample, the analysis must control analytically for the individual-level pre-intervention characteristic(s) on which the groups differ (see Statistical and Analytical Issues below).

Studies that present baseline differences greater than 0.5 standard deviation on student SES also fail to meet standards. Studies that demonstrate equivalence on a pre-test measure but do not provide information on SES at baseline should be given the benefit of the doubt that the two groups are equivalent. Reviewers should note in their SRGs and in the Study Tracker Additional Notes section that baseline data on SES were not provided and assumed equivalent.

In addition, if there is evidence that the populations were drawn from very different settings (such as rural versus urban, or high-SES versus low-SES), the Chair or Evidence Coordinator may decide that the environments are too dissimilar to provide an adequate comparison.

Study Ratings

Relevant to: Full Review

Protocol:

Meets Standards—Follow WWC version 2.0 standards.

Meets Standards with Reservations—Follow WWC version 2.0 standards.

Uncertain—Follow WWC version 2.0 standards. If missing information *prevents reviewers from determining how a study should be rated (without an author query)*, the study should be marked as “Uncertain”. If other information is missing that would enhance the study description in the SRG, but would not affect the rating of the study, the “Uncertain” label should not be applied.

Does Not Meet Standards—Follow WWC version 2.0 standards.

Does Not Pass Screens—Follow this review protocol with respect to timeframe, age/grade range, location of study, intervention content or type, outcomes measured, and study design.

Study for Support, Not Review—Apply this rating to studies with a design that is not eligible for review, but that have been recommended by a panelist (for example, key literature reviews or meta-analyses in the field). These studies will be assigned to staff (not necessarily reviewers) who will summarize them and draw out key messages that are relevant to the guide. Some studies not passing the screen, but related to a key point in the guide and providing a low level of evidence, may be assigned this rating by the Evidence Coordinator after the study is screened out of the review process.

Missing Information

Relevant to: Full Review

Protocol: **No author queries.** Because of the tight timeline for practice guides, authors are not contacted as part of the review process.

Any missing information should be noted on the SRG and in the Study Tracker.

Vague descriptions of study design: Studies may describe using intervention and comparison groups, but may not specify whether units were randomly assigned. If the study does not specifically describe assignment as random, reviewers will treat the study as a QED.

Vague descriptions of group formation: Studies may describe randomization of intervention and comparison groups, but may not specify the number of units in each group. If this is the case, practice guide reviewers should contact Shannon Monahan or Jeffrey Max—these instances will be handled on a case-by-case basis.

Mismatch between unit of assignment and unit of analysis. In some studies, the unit of assignment may not match the unit of analysis. For example, the researchers may assign teachers to conditions, but may assess the outcomes on the students of those teachers. In addition to following WWC procedures for cluster adjustments in these mismatch cases, reviewers should take special care to note how analysis groups were formed when the units of assignment were randomly assigned (for example, teachers were randomly assigned to condition, and five students from each classroom were analyzed). To preserve random assignment, participants must be assigned to condition based only on chance, and each participant must have a nonzero probability of being assigned to a condition. A study randomly assigning teachers and administering the intervention to five students selected prior to randomization or by an empirical method (for example, those who performed lowest on a standardized assessment) is considered an RCT because the principles of random assignment are preserved. But, a study randomly assigning teachers and asking the teachers to subsequently recommend or select (in some subjective way) the students in their class to participate in the study is considered a QED and must demonstrate baseline equivalence of the analysis sample of students.

Overalignment of measures. If reviewers cannot determine whether outcome measures are overaligned with the intervention, the study will be rated “uncertain”. The SRG and Study Tracker notes field should clearly describe that this is the reason for uncertainty. The Panel Chair will assess the measures used in the study and provide guidance to the reviewers.

Statistical and Analytical Issues

Relevant to: Full Review

Protocol:

RCT studies with low attrition do not need to use statistical controls in the analysis, although statistical adjustment for well-implemented RCTs is permissible and can help generate more precise effect size estimates. For RCTs, the effect size estimates will be adjusted for differences in pre-intervention characteristics at baseline (if available) using a difference-in-differences method if the authors did not adjust for pretest (see Appendix B of the *Handbook*). Beyond the pre-intervention characteristics required by the equivalence standard, statistical adjustment can be made for other measures in the analysis as well, though they are not required.

For the WWC review, the preference is to report on and calculate effect sizes for post-intervention means adjusted for the pre-intervention measure for RCTs and QEDs. If a study reports both unadjusted and adjusted post-intervention means, the WWC review will report the adjusted means and unadjusted standard deviations. If adjusted post-intervention means are not reported, the WWC reviewers will use the difference-in-differences procedure.

The statistical significance of group differences will be recalculated if (a) the study authors did not calculate statistical significance, (b) the study authors did not account for clustering when there is a mismatch between the unit of assignment and unit of analysis, or (c) the study authors did not account for multiple comparisons when appropriate. Otherwise, the review team will accept the calculations provided in the study.

When a misaligned analysis is reported (that is, the unit of analysis in the study is not the same as the unit of assignment), the effect sizes computed by the WWC will incorporate a statistical adjustment for clustering. The default intraclass correlation used for the Fractions Practice Guide review is 0.20 for math outcomes. For an explanation about the clustering correction, see Appendix C of the *Handbook*.

When multiple comparisons are made (that is, multiple outcome measures are assessed within an outcome domain in one study) and not accounted for by the authors, the WWC accounts for this multiplicity by adjusting the reported statistical significance of the effect using the Benjamini-Hochberg correction. See Appendix D of the *Handbook* for the formulas the WWC uses to adjust for multiple comparisons.

All standards apply to overall findings as well as analyses of sub-samples.

Description of the Intervention

Relevant to: Full Review

Protocol:

Definition of “Intervention.” The guide will consider studies of branded comprehensive or supplemental curricula or effective and replicable strategies for teaching fractions to students in kindergarten through 8th grade. These may include strategies or curricula used by teachers in classrooms, those used by math specialists in the school, or those for use by paraprofessional educators, tutors, or parents.

Relationship of Interventions to the Panel Recommendations. The proposed practice guide will examine practices to improve conceptual understanding of fractions, operational fluency with fractions, and problem-solving skills with fractions.

Recommendations in the guide will center on questions like:

- What strategies are effective for teaching fraction concepts and building fluency with fraction computations?
- How can teachers improve students’ conceptual knowledge and procedural knowledge of fractions?
- How can teachers build on students’ intuitive sense of fraction concepts and their “pre-fraction knowledge” to teach fractions?
- What role should different representations of fractions (for example, number line, pictorial representations) play in the teaching of fractions?
- What instructional strategies can be used to address the common challenges students face in learning fractions?
- What type of knowledge and skills are needed to effectively teach fractions?

Recommendations will NOT address the following: professional development, teacher preparation, and textbook design issues.

Criteria for SRG Description. A thorough description of the practices or curricula each intervention and/or comparison group received is required to assist the panelists with successfully incorporating the reviewed evidence into the guide. However, no study should fail to meet standards due to an incomplete description. Reviewer summaries of intervention descriptions (for inclusion in the study tracker) should include the following factors, if available:

- Types and number of activities
- Grade level or age
- Setting (school and classroom)
- Lesson duration and number of lessons
- Time between pretest and posttest
- Information on the fidelity of implementation
- Staff who delivered the intervention (teacher, math specialist, researcher, etc.)
- Training of staff delivering the intervention