REVIEW PROTOCOL FOR PRACTICE GUIDE
ASSISTING STUDENTS STRUGGLING WITH MATHEMATICS: INTERVENTION IN THE ELEMENTARY AND MIDDLE SCHOOL GRADES
Version 4.0 (December 2018)

This review protocol guides the review of research that informs the What Works Clearinghouse (WWC) *Assisting Students Struggling with Mathematics: Intervention in the Elementary and Middle School Grades* practice guide (from here on referred to in this document as mathematics intervention practice guide). The review protocol is aligned with the WWC Procedures and Standards Handbooks Version 4.0.

PURPOSE STATEMENT

The mathematics intervention practice guide will be an update of the original practice guide, *Assisting Students Struggling with Mathematics: Response to Intervention (RtI) for Elementary and Middle Schools*, published in 2009. This update is in response to the nature and quantity of research in the field and an increased interest in interventions for students struggling in mathematics.

Over the last 10 years, there has been an increase in rigorous research in Grades K–8 addressing interventions that could be used in Multi-tiered Systems of Support (MTSS). The intervention research has focused on a range of mathematical topics (e.g., fractions, ratios, proportional reasoning). The updated practice guide, like the original guide, will focus on the most effective instructional practices in mathematics for the target population—students in kindergarten through Grade 8 who struggle learning mathematics (i.e., students with mathematics difficulties [MD] or students with learning disabilities in mathematics [MLD]).

The following research questions guide the evidence review effort for the updated mathematics intervention practice guide:

- Which instructional practices or approaches recur in effective interventions for students in Grades K–8 requiring intervention in mathematics?
- Are there effective intervention practices that impact student understanding and proficiency in any of the following topic areas?
    1) Counting and cardinality
    2) Whole numbers
    3) Rational numbers
    4) Algebra and algebraic reasoning
    5) Geometry
    6) Statistics

KEY DEFINITIONS

Mathematics Intervention. In this review, a mathematics intervention is defined as a replicable individual (1:1), small-group (2 to 6 students), or large-group (more than 6 students)[1] intervention aimed at helping students with disabilities or those at risk for learning mathematics. The intervention need not be a part of a fully developed Response to Intervention (RtI) or Multi-tiered Systems of Support.

Multi-tiered Systems of Support (MTSS). MTSS is a framework for providing instructional services to all students and additional supports to students who struggle. This framework typically includes three tiers of instruction—Tier 1: core mathematics classroom instruction, Tier 2: small-group/large-group intervention, and Tier 3: intensive intervention. Some schools and districts conceptualize MTSS with additional tiers of instruction and may include up to 5 tiers. This term is often used interchangeably with Response to Intervention (RtI).

Mathematical Learning Disability (MLD). Students with a learning disability in mathematics demonstrate a specific impairment in mathematics. As the standards for determining the presence or absence of MLD vary and there is disagreement concerning the definition, this review will include all classifications made according to the policy of the study's state or district. For the purposes of this review, the terms *math learning disorder*, *students with a learning disability in math*, and *dyscalculi*a will be considered as synonymous and represented by MLD.

Mathematics Difficulties (MD). Students with mathematics difficulties who fall below grade level in mathematics generally, or in a specific topic area (e.g., counting, fractions, positive or negative integers). This at-risk status is usually determined by assessment data. Specific criteria for determining the at-risk status of a student are described in the Eligible Student Populations section. For the purposes of this review, the terms *mathematical difficulties*, *at-risk*, *low achieving*, and *poor math achievement* will be considered as synonymous.

ELIGIBILITY CRITERIA

Eligible Populations

- ***Location.*** The intervention must be provided to students in an academic setting, including before school, after school, or in summer school programs.
- ***Grade range.*** Students with a learning difficulty or disability in mathematics or those at risk for mathematics difficulties should be from kindergarten through eighth grade (ages 5 years 0 months, through 14 years 11 months), or in any subset of these grades. Studies that contain students in other grades will not be included unless (a) study reports disaggregated results for students in eligible grades, or (b) students in eligible grades represent the majority of the aggregated mixed-age sample. If the study does not make explicit the number of students in each grade, a study will be included if 50% or more of the grades included in the sample fall within the eligible grade range. If the study provides only the

[1] For the purpose of this review, a large group with as many as 20 students can be considered an intervention if it is offered above and beyond a student's core mathematics instruction. At the middle school level, these intervention classes may be called "double dose," "elective mathematics," or "foundational mathematics."

mean age of the sample without any grade information, the mean age must be larger than 5 years 0 months but smaller than 14 years 11 months.

- ***Students*.** At least 50% of the students in a study must include:
  1) Students formally classified as having a learning disability in mathematics; and/or
  2) Students identified as being at risk for failure with grade-level mathematics content (i.e., students with mathematics difficulties). The at-risk status can be met through any one of the following criteria:
     a) Students performing below the 40th percentile on a nationally or locally norm-referenced measure of mathematics;
     b) Students performing below proficient on state mathematics assessments;
     c) Students whose score on a valid screening measure of mathematics knowledge indicates they are performing at least two grade levels below the criterion for their grade or are likely to perform at the lowest one-third of their class, school, or district;
     d) Students whose score on a relevant pretest with face validity indicates that they are performing below average on the skills being targeted by the intervention; or
     e) Students identified by their teacher as having reading difficulties, with supporting data providing evidence of those difficulties.

Note: Teacher nomination by itself is not sufficient and needs to be corroborated with data.

When using criteria (b), Principal Investigator (PI) input is needed. If a sample includes a wide range of student ability levels, then the study will be eligible for inclusion if a sub-analysis is conducted for the population of interest. This review does not have any subgroups of interest.

Eligible Mathematics Interventions

The review will consider studies of intervention programs or approaches or sets of instructional practices for teaching the target population (i.e., students with a learning disability in mathematics or students at risk for mathematics difficulties in kindergarten through eighth grade).

The intervention could be a small-group (2 to 6 students) or large-group (more than 6 students) intervention, or an intervention provided in a one-on-one setting. Technology-based or blended interventions are also eligible. The mathematics interventions may be "bundled" interventions consisting of a set of practices (e.g., use of think-alouds and modeling, feedback that includes precise mathematical language, and use of manipulatives).

These intervention programs or instructional practices could be implemented by a variety of instructional personnel (e.g., general education teachers, special education teachers, paraprofessionals, trained graduate students, or trained volunteers).

Only mathematics interventions that are replicable are eligible for review. The following characteristics of an intervention must be known to reliably reproduce the intervention with different participants, in other settings, and at other times:

- Skills or proficiencies being targeted, approach to enhancing the skill(s) and proficiencies (e.g., strategies, activities, and materials);
- Unit of delivery of the intervention (e.g., whole group, individual);
- Medium/media of delivery (e.g., teacher-led instruction or software);
- Intervention duration and intensity (i.e., group size, intervention length);
- Description of individuals delivering or administering the intervention; and
- Description of the students receiving the intervention.

Both "branded" and "non-branded" interventions will be reviewed. Branded interventions are commercial or published programs and products that may possess any of the following characteristics:

- An external developer who provides technical assistance (e.g., instructions/guidance on the implementation of the intervention) or sells or distributes the intervention; or
- Trademark or copyright.

The review excludes (a) practices related to professional development or teacher preparation, and (b) other interventions not appropriate for a practice guide on instruction for students struggling in mathematics, such as comprehensive school reform.

Eligible Research

The *WWC Procedures Handbook Version 4.0* discusses the types of research reviewed by the WWC in Section II: Developing the Review Protocol and Section III: Identifying Relevant Literature. Master's theses and dissertations are ineligible for review. However, dissertations will be used as supplemental material if the study has been published in a peer-reviewed journal and is eligible for review. In this review, the following additional parameters define the scope of research studies to be included:

- ***Topic.*** The study must focus on the effects of a mathematics intervention curricula or replicable instructional practices on one or more outcome measures of mathematics achievement. Mathematics interventions that are eligible for this review are described in the Eligible Mathematics Interventions section. Relevant domains for the outcome measures are described in Eligible Outcome Domains section.
- ***Time frame.*** The study publication date must be between 2004 and 2018. Studies must be publicly available (accessible online or available through a publication, such as a journal) at the time of the original or updated literature search. Any studies from the original practice guide that met WWC design standards and fit the current set of recommendations generated by the panel will also be reviewed using WWC Group Design Standards Version 4.0, even if they were published prior to 2004.
- ***Sample.*** The study sample must meet the requirements described in the Eligible Populations section.

- ***Language.*** The study must be available in English to be included in the review. Studies examining mathematics interventions in languages other than English will be included in the review as long as the study report is in English.
- ***Location.*** The study must be conducted in the United States, its territories or tribal entities, or at U.S. military bases overseas.

Eligible Outcome Domains

The outcome measures for assessing impacts will need to be from the following domains:

Domain name	Description
Counting and Cardinality	Includes measures that assess number recognition, students' ability to count, understand one-to-one correspondence, and determine the total number of items in a set.
Whole Numbers Computation	Includes measures that assess students' ability to add, subtract, divide, or multiply whole numbers.
Whole Numbers Word Problems/Problem Solving	Includes measures that assess students' ability to solve word problems (including non-routine problems) that require use of addition, subtraction, division, and/or multiplication of whole numbers.
Whole Numbers Magnitude Understanding/Relative Magnitude Understanding	Includes measures that assess students' understanding of place value or their ability to estimate the magnitude of a whole number, compare the magnitude of two whole numbers, or locate whole numbers on a number line.
Whole Numbers Knowledge	Includes measures that assess achievement across two or more eligible outcome domains for whole numbers described above (i.e., a combination of word problems, magnitude, and/or computation).
Rational Numbers Computation	Includes measures that assess students' ability to add, subtract, divide, or multiply rational numbers (fractions and/or decimals).
Rational Numbers Word Problems/Problem Solving	Includes measures that assess students' ability to solve word problems (including non-routine problems) that involve use of rational numbers (fractions, ratios, proportions, and/or decimals).
Rational Numbers Magnitude Understanding/Relative Magnitude Understanding	Includes measures that assess students' ability to estimate the magnitude of a rational number or compare the magnitude of two rational numbers (fractions and/or decimals).
Rational Numbers Knowledge	Includes measures that assess achievement across two or more eligible outcome domains for rational numbers described above (i.e., a combination of word problems, magnitude, and/or computation).

Domain name	Description
Algebra and Algebraic Reasoning	Includes measures involving operations and concepts in algebra (e.g., understanding the equal sign, patterns, functions, positive and negative integers, linear equations). Measures may also include non-conventional or non-standard equations (i.e., the unknown is not always on the right-hand side of the equal sign).
Geometry	Includes measures involving operations and concepts in the area of geometry (e.g., area, congruence, volume).
Statistics	Includes measures involving operations and concepts in the area of statistics (e.g., probability, making inferences and predictions based on data, and data interpretation).
General Mathematics Achievement	Includes measures that assess mathematics achievement across content areas (e.g., rational number computation and whole number computation; algebra and geometry), or that combine an eligible outcome domain described above with another mathematics outcome domain that is not eligible for review under this protocol.

The outcome domains for each recommendation will be limited to only those that are considered *relevant* by the panel for that recommendation. This determination will be made by the panel after the recommendations have been identified.

Relevant outcome measures include a wide array of measures, ranging from nationally normed achievement tests, state or local tests of mathematics achievement, and researcher-developed measures that assess students' mathematical understanding, procedural capability, and problem-solving ability in the above domains.

Other, more discretionary measures of mathematics performance, such as student grades assigned by teachers, are not eligible for review.

Measures administered after the completion of the intervention are acceptable outcome measures for this guide. To consistently examine effects across different interventions, measures administered closest to the end of the intervention will be considered as the primary posttest and will be used to determine the level of evidence. All other outcomes (e.g., delayed or follow-up posttest, transfer measures) will also be included in the SRG as they will be reported in the practice guide as supplementary outcomes.

EVIDENCE STANDARDS

Eligible studies will be reviewed using the design standards described in the *WWC Standards Handbook Version 4.0*. Only issues that are unique to this review are discussed below. For all other relevant considerations (e.g., reliability of outcome measures, statistical adjustments), refer to the *WWC Standards Handbook Version 4.0*.

Sample Attrition

This review uses the *optimistic* boundary for attrition. This boundary is selected based on the assumption that most attrition in the studies on interventions for students struggling in mathematics is likely due to factors that are not strongly related to intervention status. For example, these factors may include family mobility or absences on days that assessments are given.

Joiners in Cluster Randomized Controlled Trials (RCTs)

The WWC defines a *joiner* as any student who enters a cluster after the results of random assignment are known to any individual who could plausibly influence a student's placement into a cluster (for example, parents, students, teachers, principals, or other school staff). The presence of joiners in an analytic sample has the potential to introduce bias into estimates of an intervention's effectiveness.

For this review, in cluster RCTs where the unit of assignment is a group (or classroom), *all joiners pose a risk of bias*. This is because students might be assigned to groups based on knowledge of the intervention. Additionally, students or parents may influence their assignment to clusters (e.g., classrooms) because they may have a specific preference for or against the intervention. Therefore, a study that includes at least one such joiner in the analytic sample has a risk of bias from joiners.

In cluster RCTs where the unit of assignment is a school or a group of schools (such as a district), *no joiners pose a risk*, as the intervention is unlikely to influence school enrollment or placement decisions. However, the PI and the review team leadership has discretion to revise this assessment.

Additionally, when an intervention and unit of assignment in a cluster RCT do not fall into the two categories described above, the PI and the review team leadership have discretion to make a decision on whether the joiners pose a risk of bias.

Baseline Equivalence

If the study design is an RCT or regression discontinuity design (RDD) with high levels of attrition or a quasi-experimental design (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.

Baseline Equivalence of Individuals

For studies that must satisfy baseline equivalence of individuals, including cluster-level assignment studies being reviewed for evidence of effects on individuals, baseline equivalence needs to be established for the analytic intervention and comparison groups using:

- A pre-intervention measure of the outcome used in the analysis; or
- If a pre-intervention measure of the outcome used in the analysis is not available, a pre-intervention measure of an outcome from any of the 10 outcome domains detailed in the Eligible Outcomes section can be used. For example, a pretest from the general

mathematics achievement domain can be used to establish baseline equivalence for an algebra outcome when an algebra pretest is unavailable.

The baseline equivalence will be assessed for each analytic sample on an outcome-by-outcome basis. Specifically, baseline equivalence for an eligible outcome measure will be assessed based on the magnitude of baseline differences for its most closely associated pre-intervention measure. Baseline differences on other pre-intervention measures do not influence the assessment of baseline equivalence for the outcome measure. For example, if both pre- and post-intervention measures of outcomes A, B, and C are available and the baseline difference for the pre-intervention measure of C is outside of the statistical adjustment range (that is, it exceeds 0.25 standard deviations), then the finding for this outcome would be rated *Does Not Meet WWC Design Standards* because it does not satisfy the baseline equivalence requirement. A finding for outcome B, however, could still meet design standards if the authors satisfy the baseline equivalence requirement using the pre-intervention measure of B. The same is true for outcome A. In addition, when the baseline difference for a pre-intervention measure is in the statistical adjustment range (that is, it is between 0.05 and 0.25 standard deviations), the adjustment must be made only in the analysis of the associated outcome measure. For example, if the pre-intervention difference in B requires statistical adjustment, then only the analysis of outcome B must be adjusted.

For this review, it is not necessary to demonstrate equivalence on student, teacher, or school demographic characteristics.

Baseline Equivalence of Clusters

In general, considerations for satisfying baseline equivalence of individuals also apply to satisfying baseline equivalence of clusters. In particular, baseline equivalence of clusters in the intervention and comparison groups must be satisfied by one of the same baseline measures described above for assessing baseline equivalence of individuals, and the same statistical adjustment requirements apply.

The baseline equivalence requirement for the analytic sample of clusters can also be met using data from an earlier assessment of the same cohort of individuals in the analytic sample within the same clusters. For example, if the sample includes Grade 4 students in the 2015–2016 school year, the baseline requirement can be met using end-of-year data for the same cohort in Grade 3 from the 2014–2015 school year.

Outcome Measure Requirements

The *WWC Standards Handbook Version 4.0* discusses the types of outcomes, criteria the outcome must meet, and how outcomes are reported in Section IV.A: Outcome Requirements and Reporting. In particular, this review follows the requirements stated in the *Standards Handbook Version 4.0* regarding the reliability of outcome measures.

Statistical Adjustments

The *WWC Procedures Handbook Version 4.0* discusses the types of adjustments made by the WWC in Section VI: Reporting on Findings. For "mismatched" analysis (i.e., when a study assigns units at the cluster level but conducts analysis at the individual level), this topic area uses the WWC

default intra-class correlation coefficient for achievement outcomes of 0.20 for all eligible outcomes, unless a study-reported intra-class correlation coefficient is available.

Eligible Study Designs

Studies that use group designs (RCTs and QEDs) or RDDs are eligible for review using the appropriate standards or pilot standards. Single-case designs (SCDs) are not eligible for review.

PROCEDURES FOR CONDUCTING THE LITERATURE SEARCH

The *WWC Procedures Handbook Version 4.0* discusses the procedures for conducting a literature search in Section III: Identifying Relevant Literature and Appendix B: Policies for Searching Studies for Review. The following search terms will be used to conduct the literature search.

Keywords	Related Search Terms
Intervention	Intervention, RTI, response to intervention, response to intervention instruction, multi* tier*, system of support, MTSS, remedia*
Content	Fraction*, whole number, rational number, numeracy, decimal, algebra*, geometry, statistics, math*, arithmetic, number knowledge, number understanding, computation, word problem, problem solving
Population	Kindergarten, K, first grade, 1st grade, grade 1, second grade, 2nd grade, grade 2, third grade, 3rd grade, grade 3, fourth grade, 4th grade, grade 4, fifth grade, 5th grade, grade 5, sixth grade, 6th grade, grade 6, seventh grade, 7th grade, grade 7, eighth grade, 8th grade, grade 8, elementary grades, upper elementary grades, middle school, junior high AND At risk, struggl*, math* disab*, math* diff*, learning disab*, low* achiev*
Study Design	ABAB design, alternating treatment, experiment*, QED, quasi-experiment*, quasiexperiment*, random assignment, RCT, randomized controlled trial, RDD, regression discontinuity, simultaneous treatment, SCD, single case, single subject, multiple baseline, changing criterion, intrasubject replication, reversal design, withdrawal design