

**REVIEW PROTOCOL FOR PRACTICE GUIDE
ASSISTING STUDENTS STRUGGLING WITH READING: INTERVENTION IN GRADES 4-9
Version 4.0 (January 2020)**

This review protocol guides the review of research that informs the What Works Clearinghouse (WWC) *Assisting Students Struggling with Reading: Intervention in Grade 4-9* practice guide. The review protocol is aligned with the [WWC Procedures and Standards Handbooks Version 4.0](#).

PURPOSE STATEMENT

The latest version of the National Assessment of Educational Progress (NAEP, 2019) found that 34 percent of grade 4 students and 27 percent of grade 8 students are performing below the basic level in reading. The NAEP results also show a decline in the percentage of students performing at the basic level in reading from previous years.

Until recently, the majority of reading intervention research has focused on the early elementary grades, with the hope that successful early interventions would lead to sustained gains in later grades. Over the past decade, there has been an increase in the number of studies that focus on interventions for struggling readers in grades 4 and 5, as well as in middle school and freshman year of high school.

As struggling readers move up the grades, they continue to experience difficulties with fluent word, sentence, and paragraph reading; academic vocabulary; and/or comprehension, which in turn severely impact their access to informational and narrative text across content areas. Providing recommendations on evidence-based practices for effective reading intervention for students in grades 4-9 seems equally important as for students in grades K-3. This practice guide, *Assisting Students Struggling with Reading: Intervention in Grades 4-9*, addresses the need to provide evidence-based practices about interventions and supports for struggling readers in grades 4-9.

KEY DEFINITIONS

Reading Intervention. For this review, a reading intervention is defined as a curriculum or set of practices aimed at helping students with reading disabilities or those with reading difficulties. The intervention need not be a part of a fully developed Response to Intervention (RtI) or Multi-tiered System of Support (MTSS). However, the interventions are those that are implemented above and beyond the Tier 1 core instruction. Reading courses provided in middle school or high school qualify for review.

Reading Disability. Students with reading disabilities demonstrate a specific impairment in learning to read. Since the standards for determining the presence or absence of a reading disability 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 by deferring to the classification reported by the authors. For the purposes of this review, the terms *reading disability*, *dyslexia*, or *learning disability in reading* will be considered synonymous.

Reading Difficulties. Students with reading difficulties are those whose scores on a standardized test or a valid screener indicate that their reading performance is below proficient or below grade level, or that they are at risk of falling behind without intervention. These difficulties may be evident in only some of the domains of reading (e.g., weak decoding and word recognition skills; poor comprehension of informational text). 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 *reading difficulties*, *at-risk*, *low achieving*, and *poor reading achievement* will be considered synonymous.

Struggling Readers. For the purposes of this review, the term struggling readers refers to students with reading disabilities as well as students with reading difficulties.

ELIGIBILITY CRITERIA

Eligible Population

Grade range. Students with reading disabilities or reading difficulties should be from grade 4 through grade 9 (typically between ages 9 and 15). Studies that contain students in other grades will not be included unless (a) the 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 only if 50% or more of the grades included in the sample fall within the eligible grade range. If the study provides only the mean age of the sample without any grade information, the mean age must be larger than 9 years 0 months but smaller than 15 years 11 months.

Students. At least 50% of the students in a study must include:

1. Students formally classified as having a reading disability; and/or
2. Students identified as being at risk for failure in reading (i.e., students with reading difficulties). The at-risk status can be met through any one of the following criteria:
 - a. Students performing below the 35th percentile on a nationally or locally norm-referenced measure of reading; or
 - b. Students performing below proficient on state assessments; or
 - c. Students performing two or more years below grade level; or
 - d. Students whose score on a valid reading screening measure indicates that they are likely to perform in the lowest 35 percent of their class, grade level, school, or district; or
 - e. Students identified by their teacher as having reading difficulties, with supporting data providing evidence of those difficulties.

When using criteria (d) or (e), input from the Principal Investigator (PI) of the practice guide is needed.

Eligible Subpopulation

This review will consider outcomes reported for a sub-sample of English learners.

Eligible Reading Interventions

This review will consider studies of curricula or sets of instructional practices for teaching the target population (i.e., students with reading disabilities or students with reading difficulties in grade 4 through grade 9). This will include instruction or intervention provided to the target students over and beyond typical core classroom Tier 1 instruction.

The reading intervention can be intended for any number of students. The intervention need not be conducted in a school setting and can be implemented before or after school, during the summer or on weekends.

The reading intervention could be implemented by a variety of instructional personnel (e.g., teachers, special educators, paraprofessionals, volunteers, or trained project staff).

The reading intervention may be a “bundled” set of instructional practices (e.g., use of think-alouds and modeling, feedback that guides students to correct responses, strategy instruction for improving comprehension, vocabulary instruction). The intervention could also be a technology-based or a blended online and in-person program.

Only reading interventions that are replicable are eligible for review. The following characteristics of an intervention must be documented, so that others can reliably reproduce the intervention with different participants, in other settings, and at other times:

- Domain(s) of reading being targeted;
- Curriculum or sets of instructional practices for teaching the target population;
- Intervention group size (e.g., small group, large group, individual);
- Mode of delivery of instruction (e.g., instruction organized and delivered by an adult, instruction delivered by software, or “blended intervention” that contains both types of instruction);
- Who administers the intervention (e.g., teachers, special educators, paraprofessionals, volunteers, or trained project staff);
- Intervention duration (number of sessions and length of typical session); and
- Description of the students receiving the intervention.

Both “branded” and “non-branded” interventions will be reviewed. Branded interventions are published programs that could be distributed by a commercial firm or a research organization. These interventions may have a trademark or copyright.

This review excludes studies solely focused on professional development in the absence of instructional intervention.

Eligible Studies

Time frame. The study publication date (print or online) must be between January 1, 2005 and March 31, 2020. Studies must be publicly available and accessible online.

Language. The study must be available in English, and the focus of the intervention must be on improving reading in English.

Location. The study must be conducted in the United States, its territories or tribal entities, or at U.S. military bases overseas.

Designs. The study must include a randomized controlled trial (RCT), a quasi-experimental design (QED), a regression discontinuity design (RDD), or single case design (SCD).

Eligible Outcome Domains

Eligible domains are listed below. The PI along with the panel will determine which domains are relevant for each recommendation.

Domain Name	Description
Word and pseudoword reading	Includes measures that assess students' ability to decode real and pseudo words (i.e., nonsense words), including measures of accuracy or fluency.
Passage reading fluency-oral	Includes measures that assess students' ability to read connected text orally.
Passage reading fluency-silent	Includes measures that assess students' ability to read connected text silently.
Reading vocabulary	Includes measures that assess students' ability to demonstrate knowledge of words that they read.
Listening comprehension	Includes measures that assess students' ability to respond to comprehension questions after listening to a passage or story that has been read to them, or to provide an accurate retelling of the passage.
Reading comprehension	Includes measures that assess students' ability to answer questions in passages that address factual information, inferencing, drawing conclusions, identifying main ideas, and determining the meaning of words from context, or to provide an accurate retelling of the passage.
Measures of general reading proficiency and English Language Arts	These typically include state assessments and nationally normed tests. These could also include measures that assess achievement from more than one outcome domain (e.g., both reading vocabulary and reading comprehension).

Eligible Measures

Only measures from the eligible outcome domains can be included in this review. The measures can be researcher-developed or standardized assessments.

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 main posttest and be used to determine the level of evidence.

The following outcomes will be included as supplemental outcomes: (a) posttest measures administered closest to the end of the intervention for the eligible subpopulation of English learners, (b) all delayed or follow-up posttest measures, and (c) outcome measures described as transfer measures.

Discretionary measures of reading performance, such as student grades assigned by teachers, are not eligible.

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 of interventions for struggling readers 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) within a school, *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.

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 decide whether the joiners pose a risk of bias.

Baseline Equivalence. If the study design is an RCT with high levels of attrition, an RDD 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.

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, reviewers can use a pre-intervention measure of an outcome from any of the outcome domains detailed in the Eligible Outcome Domains section, except listening comprehension. For example, a pretest from the general reading proficiency domain can be used to establish baseline equivalence for a fluency outcome when a fluency pretest is unavailable.

The baseline equivalence will be assessed for each analytic sample on an outcome-by-outcome basis. Baseline equivalence for an eligible outcome measure will be assessed for its most closely associated pre-intervention 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 exceeds 0.25 standard deviations, then the finding for outcome C would be rated *Does Not Meet WWC Design Standards*. 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, between 0.05 and 0.25 standard deviations, the adjustment is only made for 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. Baseline equivalence of clusters in the intervention and comparison groups must be satisfied by one of the same baseline measures 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 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 outcomes 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.

PROCEDURES FOR CONDUCTING THE LITERATURE SEARCH

The literature search will be conducted using the public ERIC search engine (<https://eric.ed.gov/>) which affords relevant results by automatically searching for descriptors and their variants across records without advanced logic or Boolean operators. Therefore, the following two search terms will be used to conduct two comprehensive literature searches using ERIC: *reading intervention* and *reading tutoring*. The search results will be further limited by the descriptor *reading difficulties*. The remaining records will be screened for eligible topics, populations, and study designs prior to review.