

WWC EVIDENCE REVIEW PROTOCOL FOR MIDDLE SCHOOL MATHEMATICS INTERVENTIONS, VERSION 1.0

Topic Area Focus

The What Works Clearinghouse (WWC) review focuses on curriculum-based interventions, which contain learning goals that spell out the mathematics that students should know and be able to do, instructional programs and materials that organize the mathematical content, and assessments. There are two WWC Evidence Reports planned in this topic area. The first report focuses on interventions for middle school mathematics; the second looks at interventions for elementary school mathematics.

This first report uses systematic reviews of evidence to address the following questions:

- Which curriculum-based interventions are effective in increasing the learning of mathematics content and skills (that is, what students should know and be able to do) among middle school students?
- Are some interventions more effective for certain types of students, particularly students who lag behind in mathematics achievement?

Key Definitions

Curriculum-based interventions. A curriculum-based math intervention will be defined in this review as a replicable, materials-based instructional program central to students' regular learning activities for which

- the population of learners is specified as middle school students;
- the learning goals for students are well delineated;
- student outcomes related to mathematics achievement are directly assessed;
- one or more of the following content areas are covered: number and operations, algebra, geometry, measurement, data analysis and probability; and
- instruction extends over the course of one semester or more.

Curricula may be based on text materials, computer software, videotapes, or any other materials base, or combination thereof. Interventions of short duration (e.g., less than one semester) or that merely supplement an existing curriculum (e.g., tutorial program) are excluded.

We will consider only core curriculum for continuous review. Core math curricula are defined as instructional programs that

- are comprehensive;
- extend over the course of one semester or more;
- are central to students' regular school instruction; and
- are based on text materials, manipulatives, computer software, videotapes, other materials, or any combination thereof.

Mathematics Content and Skills

Mathematics content and skills are commonly described as what students should know and be able to do. Mathematics content varies somewhat across curricula and grade level, but generally includes: number and operations, algebra, geometry, measurement, and data analysis and probability. Mathematics skills are the application of the learning of this content, as well as an understanding of mathematical concepts, procedures, and problem solving. These include: problem solving, reasoning and proof, making connections, oral and written communication, and uses of mathematical representation.

GENERAL INCLUSION CRITERIA

Populations to be Included

The WWC reviews interventions that are designed for middle school students. However, because there is some variation across school districts around the nation on how middle school is defined and organized, for this review middle school is primarily defined as any school with any of the four grades from grades 6 through 9. This definition of middle school is used throughout the protocol. However, a working definition is needed since some school districts might still be using the traditional junior high school building configuration where 7th and 8th grades are isolated from 6th and 9th grades. The first criterion in screening is to include studies that explicitly indicate the schools were middle schools or that they were using a curriculum meant for middle schools. Otherwise, a study is included if it indicates that its sample was either 6th, 7th, 8th, or 9th grade or some combinations of these grades.

Types of Interventions to be Included

The interventions included are determined after an exhaustive search of the published and unpublished literature by the Evidence Report Team. Only research on interventions that are replicable (i.e., documented well enough that they can be reproduced) and materials-based (although the materials may come in various forms) is reviewed. Examples of possible interventions to be included are

- textbooks and textbook series,

- software programs and other educational technology that serve as the basis for well-defined curricula, and
- materials including videotapes that serve as the basis for well-defined curricula.

Types of Research Studies to be Included

This review includes empirical studies published since 1983 that focus on math interventions for students in middle schools (grades 6–9). The studies include students attending schools in the U.S. and its territories or tribal entities. Research on interventions developed outside of the U.S. is considered if the studies use U.S. samples. The focus of the outcome measures is the students, not the teachers, and at least one of the outcome measures must focus on math achievement and demonstrate sufficient reliability or face validity.

The review focuses on empirical studies, using quantitative methods and inferential statistics. Specifically these include well-conducted randomized controlled trials (RCTs), quasi-experiments with matching or equating on the pretest, and regression discontinuity designs. At this time, the WWC has not developed standards for reviewing or reporting on single-case design studies. Consequently, studies with a single-case design are not currently included in this review. The empirical focus is reflected in the collection, review, and reporting of the research. The studies must all have adequate statistical reporting such that effect sizes can be calculated.

SPECIFIC TOPIC PARAMETERS

The following parameters specify which studies are considered for analyses and which aspects of those studies are coded for the review.

1. Commonly shared and/or theoretically derived characteristics of the intervention that should be reflected in its definition and implementation include:

- Materials-based curriculum targeted to grade levels 6 through 9 or labeled as a “middle school” curriculum
- Math curriculum that covers one or more of the following:
 - Content:
 - Number and Operations
 - Algebra
 - Geometry
 - Measurement
 - Data Analysis and Probability

– Skills:

- Problem Solving
 - Reasoning and Proof
 - Making Connections
 - Oral and Written Communication
 - Uses of Mathematical Representation
- Clearly defined learning goals (e.g., the curriculum spells out what students will be able to do after its implementation)
 - Extended over the course of one semester or more of instruction
 - Enough information provided so that the intervention is replicable
 - Published between 1983 and 2007

2. The variations in the intervention that are important to examine include:

- Amount of professional development
- Degree of use of text materials, manipulatives, videotapes, software, or other instructional materials specified as part of the curriculum
- Degree of use of prescribed classroom structures, such as collaborative group work
- Amount of dosage provided to students (i.e., how often the intervention was delivered to students and for how long)
- Degree of implementation fidelity

If there is no evidence in the study that the intervention was altered when it was implemented, we assume that it was implemented as defined.

3. The important characteristics of the intervention that we need to know in order to reliably replicate it with different participants, in other settings, at other times include the following:

- The intervention is “branded.”
- The intervention is not “branded” but meets the following conditions:
 - The intervention is described in general terms.

- The duration of the intervention is described.
- The characteristics of the individuals who are expected to deliver the intervention are described.

4. Important classes of outcomes include measures of mathematics achievement. Three types of outcome measures are included:

- Standardized, nationally normed achievement tests that are appropriate for elementary students (e.g., Comprehensive Test of Basic Skills, Wide Range Achievement Test)
- Standardized state or local tests of math achievement
- Research-based or locally developed tests or instruments that assess students' mathematical concepts or skills

A study needs to include at least one type of math achievement measure that involves direct student assessment.

5. Evidence sufficient for an outcome measure to demonstrate each type of reliability includes:

Internal consistency: .60

Temporal stability: Test-retest estimates can be low even when the construct is well measured (due to real change in the participants) and high even when it is measured poorly (due to memory). If reviewers find a temporal stability estimate and it is the only reliability estimate for the measure, it should be flagged and discussed with the Senior Content Advisor (SCA) and Methodology Consultant (MC).

Inter-rater reliability: For inter-rater reliability, reviewers need to discern whether or not the estimate is corrected for chance. If reviewers find an inter-rater reliability estimate and it is the only reliability estimate for the measure, it should be flagged and discussed with the SCA and MC.

Studies must contain at least one relevant measure that has evidence of either face validity or reliability.

6. The interval of time in which studies should have been conducted to be appropriate for the Evidence Report:

Studies need to have been conducted within the past 20 years (i.e., with a publication date of 1983 or later). This is the default time interval for all WWC reviews. This time scale is wide enough to allow for a baseline of data with regard to traditional curricula, NAEP trends, and data with regard to Standards-based curricula. The rule for unpublished articles is the date the document reached its final form. However, there is recognition that there might be a lag between published articles and unpublished articles, which could cause a bias towards having

more unpublished articles. This bias is unlikely to cause a significant difference because of the lengthy time interval used. The experience with math so far indicates that authors do not keep unpublished articles dating to the early 80s. At the other end of the spectrum, some authors declined to have their studies reviewed until publication. Hence, their studies will only be reviewed at a later round.

7. The necessary characteristics that define the target population include:

A middle school curriculum is one that is taught to middle school students. However, because there is some variation across school districts around the nation on how middle school is defined and organized, for this review middle school is defined as any school with any of the four grades from grades 6 through 9.

8. The important characteristics of participants that might be related to the intervention's (i.e., curriculum's) effect that must be equated if a study does not employ random assignment include:

Broadly speaking, comparison studies are valid only if student population groups are roughly equivalent with regard to:

- Pretest of the outcome measure
- Grade level

A study employing a quasi-experimental design without at least a pretest is excluded from the review.

Though it is only necessary for groups to be equivalent with regard to pretest scores and grade level, the Evidence Report Team also assesses whether the groups are equivalent along the following dimensions:

- Gender
- Socioeconomic status
- Race
- Percentage of second-language students
- “At-risk” status
- Tracking level
- Special education
- Location
- Average class size

However it is unlikely that all such information will be provided in the studies and therefore is not a requirement of the review.

9. The characteristics of subgroups of participants that are important (a) to have variation on and (b) to test within a study to determine whether an intervention (i.e., a curriculum) is effective within these groups include:

- Gender
- Socioeconomic status
- Racial/ethnic breakdown
- Percentage of second-language students
- Percentage of bicultural students
- “At-risk” status

Note: The definition of “at risk” is provided by the study authors. We capture the author’s definition in the coding guide.

10. The characteristics of settings that are important to test within a study to determine whether an intervention is effective within these settings include:

- Location of the schools involved
- Homogenous groupings of students
- School type (public, private, religious)
- Average class size (small, medium, large)
- School SES (e.g., Title 1 school)
- School size

11. The appropriate interval for measuring the intervention’s (i.e., curriculum’s) effect relative to the end of the intervention is:

A math intervention may have an immediate effect as well as a long-term impact on students’ math achievement. Thus, measures taken at the end of an intervention, as well as measures taken any time thereafter, are included. Delayed measures taken several months or years after an intervention may be useful because they may provide strong evidence for an intervention’s effectiveness. The Evidence Report Team codes the length of the time interval between the end of an intervention and assessment of the outcome measure.

12. If students are drawn from the same local pool, which groups of individuals (e.g., students, teachers, parents, administrators, case workers) might have been able to

interfere with the fidelity of the comparison if they had known who was in the intervention and comparison groups?

The students, parents, and/or teachers in either the treatment or control group could know who was in which condition (participants are not blind to condition due to the nature of the intervention).

13. For research on this topic, differential attrition from the intervention and control groups is defined as follows:

Differential attrition is defined as being greater than 7% differential loss. If differential attrition is less than or equal to 7%, we assume that the bias associated with it is minimal. If it is greater than 7%, the burden of proof shifts, and the study authors need to convince us that differential attrition did not bias the effect size estimate. A post-attrition demonstration of group equivalence on the pretest is required. We define “post-attrition demonstration of group equivalence” as either a well-powered (.80) test of equivalence that is nonsignificant, or a standardized mean difference between groups of less than $d = .10$.

14. For research on this topic, severe overall attrition is defined as follows:

Severe overall attrition is defined as being greater than 20% loss. If overall attrition is less than or equal to 20%, we assume that the bias associated with it is minimal. If it is greater than 20%, the burden of proof shifts, and the study authors need to convince us that overall attrition did not bias the effect size estimate. A post-attrition demonstration of group equivalence on the pretest is required. We define “post-attrition demonstration of group equivalence” as either a well-powered (.80) test of equivalence that is nonsignificant, or a standardized mean difference between groups of less than $d = .10$.

15. The statistical properties of the data that are important in order to obtain an accurate estimate of an effect size are as follows:

For most statistics (including d-indexes), normal distribution and homogeneous variances are important properties. For correlations, there are no required desirable properties, and for odds-ratios there are no required desirable properties except the minimum of 5 observations per cell.

METHODOLOGY

Literature Search Strategies

The WWC Evidence Report Team employs comprehensive and systematic literature search strategies to identify the population of published and unpublished relevant studies. This section contains topic specific elements of the literature search (e.g., search terms, additional journals, and associations).

Key Word List

| | |
|-----------------------------|-------------------------|
| Intervention | Mathematical skills |
| Instruction | Mathematical concepts |
| Curriculum | Mathematics instruction |
| Curriculum based assessment | Mathematic* ability |
| Curriculum evaluation | Spatial ability |
| Mathematics | Properties mathematics |
| Number | Properties |
| Numbers | Middle schools |
| Algebra | Middle grades |
| Geometry | Sixth grade |
| Measurement | Seventh grade |
| Data analysis | Eighth grade |
| Probability | Ninth grade |
| Reasoning | Outcomes of education |
| Proof | Mathematics achievement |
| Connections | Mathematics w outcome |
| Communication | Outcomes |
| Representation | Achievement |
| Disposition | Learn |
| Attitude skills | Impact |
| Concepts | Enhance |
| Problem solving | Attainment |
| Mathematical aptitude | Evidence |
| | Assessment |

A combination of Boolean terms such as AND and OR are used with this keyword list. The librarian at AIR is consulted as to the appropriate combination to use and then tailor according to each specific electronic database.

Topic-Specific List of Journals to be Hand Searched

1. Journal for Research in Mathematics Education
2. Journal of Mathematical Behavior
3. Mathematical Thinking and Learning
4. Educational Studies in Mathematics
5. Journal of the Learning Sciences
6. Cognition and Instruction
7. Cognitive Science
8. Instructional Science

During continuous review, additional journal(s) may be added to the supplementary list when a handful of articles are identified from a particular journal that is not on this supplementary journal list.

Topic-Specific List of Organizations

1. Eisenhower Clearinghouse

Topic-Specific List of Associations

1. Association for Supervision and Curriculum Development (ASCD)
2. National Council of Supervisors of Mathematics (NCSM)
3. Psychology and Mathematics Education (PME)

Personal Contacts

The WWC Math Evidence Report Team solicits studies directly from experts in the field of education who work on math curriculum interventions. The Principal Investigator (PI) identifies these experts. We also contact experts using listservs dedicated to this topic and whose members are scholars working in this area.

After the identification of studies to be reviewed, we contact the authors of these studies to request similar materials and to ask them to “snowball” the process to colleagues whom they recommend for their work in this area.

Direct Contact with Publishers and Developers

The National Council of Teachers of Mathematics Exhibitors' List

The National Council of Teachers of Mathematics Exhibitors' List is searched for textbook publishers and software developers who have either published or developed an elementary, middle, or secondary school math curriculum. Organizations of this type may have research results of either pretests of the curriculum or evaluations of the curriculum after it has been implemented in schools that may not appear in refereed journals. We search the websites of these exhibitors to determine if they have published or developed a curriculum-based mathematics intervention. We use the guidelines developed by the PI for determining whether the products offered by the exhibitors constitute relevant mathematics interventions. The results of this search are reviewed by the PI before the exhibitors are contacted. Initially, a standardized email is sent to the relevant publishers and developers requesting submissions of pertinent research to the What Working Clearinghouse website. A follow-up letter also is mailed to organizations that do not respond to our email request.

There is a recognition that studies obtained from the publishers and developers might be biased towards a positive outcome. Articles and studies from these sources are to be flagged according to their source. However, given the few studies we have obtained so far for this math review, it is unlikely we will be able to conduct a meta-analysis and sensitivity analysis.

Note: Not all of the textbook publishers listed below are actually publishers, but they might provide information and links to textbook and curriculum series.

Textbook Publishers List

1. AAAS/Project 2061
2. Academic Systems Company
3. American Guidance Service, Inc.
4. AMME, Inc
5. Brooks/Cole Publishing Company
6. CCI Publishing (formally CORD Communications)
7. CharlesBridge Publishing
8. College Preparatory Mathematics
9. Continental Press
10. Curriculum Associates
11. Dale Seymour Publications—Pearson Learning Group
12. Developmental Studies Center
13. Didax, Inc
14. Digi-Block, LLC
15. Dimension 2000
16. Education Development Center
17. Encyclopedia Britannica
18. Evan-Moor Educational Publishing
19. Everyday Mathematics/SRA/McGraw-Hill
20. Foundation for Advancements in Science and Education (FASE)
21. Glencoe/McGraw-Hill
22. Globe Fearon Educational Publishing
23. Heinemann Educational Books
24. Holt Rinehart & Winston (part of Harcourt)
25. Houghton Mifflin (acquired DC Heath, publisher of Heath Mathematics Connection)
26. Key Curriculum Press
27. Lawrence Erlbaum Association, Inc.
28. Math Essentials
29. McDougal Littell
30. Options Publishing, Inc.
31. Pencil Point Press, Inc. (they already responded to say that they are not relevant)
32. Prentice Hall
33. Saddleback Educational, Inc.
34. Saxon Publishers, Inc.
35. Scholastic, Inc.
36. Scott Foresman/Addison Wesley/Silver Burdett Ginn
37. SRA/McGraw Hill
38. Teachers College Press
39. The Consortium for Mathematics and its Applications
40. W. H. Freeman & Company
41. William K. Bradford
42. Wright Group (acquired Creative Publications)

Software Developers List

1. Apple Computer, Inc.
2. Barnum Software
3. BestQuest Teaching Systems
4. Boxer Learning, Inc.
5. Carnegie Learning, Inc.
6. Edu2000 America, Inc.
7. Learning In Motion
8. MathMedia Educational Software, Inc.
9. Meridian Creative Group
10. New Century Education Corporation
11. Optical Data School Media
12. Riverdeep Interactive Learning
13. Steck-Vaughn
14. WestEd

Publishers and software developers are directed to submit their studies online at the WWC website. They are asked to call us if they have any concerns or need more information.

Social Science Citation Index (SSCI) will be used in the next round of reviews for identification of seminal papers.