Connected Mathematics Project (CMP)

**Intervention Description**

*Connected Mathematics Project (CMP)* is a math curriculum for students in grades 6–8. It uses interactive problems and everyday situations to explore mathematical ideas, with a goal of fostering a problem-centered, inquiry-based learning environment. At each grade level, the curriculum covers numbers, algebra, geometry/measurement, probability, and statistics.

**Research**

The What Works Clearinghouse (WWC) identified two studies of CMP that both fall within the scope of the Primary Mathematics topic area and meet WWC group design standards. No studies meet WWC group design standards without reservations; the two studies meet WWC group design standards with reservations. Together, these studies included 3,062 students in grades 6–8 in at least 23 schools in 10 locations.

The WWC considers the extent of evidence for CMP on the mathematics achievement of students in primary mathematics courses to be medium to large for the mathematics achievement domain, the only domain examined for studies reviewed under the Primary Mathematics topic area. (See the Effectiveness Summary on p. 4 for more details of effectiveness by domain.)

**Effectiveness**

*CMP* was found to have no discernible effects on mathematics achievement for students in primary mathematics courses.

**Table 1. Summary of findings**

<table>
<thead>
<tr>
<th>Outcome domain</th>
<th>Rating of effectiveness</th>
<th>Improvement index (percentile points)</th>
<th>Number of studies</th>
<th>Number of students</th>
<th>Extent of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematics achievement</td>
<td>No discernible effects</td>
<td>+2</td>
<td>2</td>
<td>3,062</td>
<td>Medium to large</td>
</tr>
</tbody>
</table>
Intervention Information

Background

CMP was developed by Michigan State University and is distributed by Pearson Education. Address: P.O. Box 6820, Chandler, AZ 85246. Email: k12customerservice@pearson.com. Web: https://connectedmath.msu.edu/. Telephone: (800) 848-9500.

Intervention details

CMP is an inquiry-based mathematics curriculum. Mathematical ideas are embedded in sequenced sets of tasks and explored in depth to help students develop a thorough understanding of mathematical concepts. Throughout the curriculum, students focus on problem-solving strategies, communicating their reasoning, offering proofs, and using representations.

CMP includes four courses: Grade 6 Mathematics, Grade 7 Mathematics, Grade 8 Mathematics, and Algebra I. Each course is organized into units, with seven to eight units per course, with units organized around a mathematical concept or set of related concepts, such as area and perimeter, or operations on fractions.

CMP seeks to promote students’ mathematical knowledge and understanding by helping them understand the connections among mathematics topic areas and between mathematics and other academic subjects. Students move flexibly between graphic, numeric, symbolic, and verbal representations to develop fluency in conceptual and procedural knowledge.

The publisher is currently selling the third edition of CMP. The developer’s website describes how the curriculum has been refined and enhanced from the first and second editions.

Cost

As of July 2016, the cost of CMP varies between $73.97 and $127.97 per student for the full curriculum, depending on the grade level and type of implementation (digital, print, or a combination of formats). Manipulatives kits cost between $221.97 and $313.47 each, and a teacher resource kit costs $524.97. More detailed cost information is available from the publisher.
Research Summary

The WWC identified 24 eligible studies that investigated the effects of CMP on the mathematics achievement of primary students. An additional 72 studies were identified but do not meet WWC eligibility criteria for review in this topic area. Citations for all 96 studies are in the References section, which begins on p. 5.

The WWC reviewed 24 eligible studies against group design standards. None of the 24 studies are randomized controlled trials that meet WWC group design standards without reservations. Two of the 24 studies use quasi-experimental designs and meet WWC group design standards with reservations. Those two studies are summarized in this report. The remaining 22 studies do not meet WWC group design standards.

Summary of studies meeting WWC group design standards without reservations

No studies of CMP met WWC group design standards without reservations.

Summary of studies meeting WWC group design standards with reservations

Cai, Wang, Moyer, Wang, and Nie (2011) examined the effect of the CMP curriculum by matching seven schools using the CMP curriculum to seven schools not using the curriculum on comparable demographic characteristics. All 14 schools were located in a single large, urban school district. Students in the intervention schools used CMP in the sixth, seventh, and eighth grades. Students in the comparison schools did not use CMP. The study involved a single cohort of students who were followed over six school years (2005–06 through 2010–11), from grade 6 to grade 11. All but the analyses of ninth and tenth grade outcomes meet WWC group design standards with reservations. The WWC based its effectiveness rating on outcomes measured at the end of eighth grade. The eighth-grade analytic sample included 303 students in the CMP group and 303 students in the comparison group in 14 schools. The study used the first edition of CMP.

Ridgway, Zawojewski, Hoover, and Lambdin (2002) conducted a study using a quasi-experimental design in which two classrooms from each of nine sites across the country used CMP. Five sites were in the Midwest, two were in the West, and two were in the East. For every two CMP classrooms, one non-CMP classroom was recruited for the study. The authors matched the CMP classrooms to non-CMP classrooms based on student ability, urbanicity, diversity in student population, and algebra or pre-algebra tracks. In five sites, there were both CMP and comparison classrooms. In the other four CMP sites, comparison classrooms were recruited from other locations. Data for the sixth- and seventh-grade analytic samples were collected in the 1994–95 school year and included 36 CMP classrooms and 18 comparison classrooms. Data for the eighth-grade sample was collected in the 1995–96 school year, and included 14 CMP classrooms and seven comparison classrooms. The WWC based its effectiveness rating on outcomes combined across students in all three grades. Although some intervention students used CMP in a previous school year, the findings from this study measure the effectiveness of receiving 1 year of the intervention because the pre-intervention measures were assessed at the beginning of same school year in which outcomes were measured. The analytic sample included a total of 2,456 students across these 75 classrooms. The study did not specify which edition of CMP was used, but based on the timing of the study, it was likely the first edition.
Effectiveness Summary

The WWC review of CMP for the Primary Mathematics topic area includes student outcomes in one domain: mathematics achievement. The findings below present the authors’ estimates and WWC-calculated estimates of the size and statistical significance of the effects of CMP on primary students. Additional comparisons are presented as supplemental findings in Appendix D. These supplemental findings do not factor into the intervention’s rating of effectiveness. For a more detailed description of the rating of effectiveness and extent of evidence criteria, see the WWC Rating Criteria on p. 23.

Summary of effectiveness for the comprehension domain

Table 3. Rating of effectiveness and extent of evidence for the mathematics achievement domain

<table>
<thead>
<tr>
<th>Rating of effectiveness</th>
<th>Criteria met</th>
</tr>
</thead>
<tbody>
<tr>
<td>No discernible effects</td>
<td>In the two studies that reported findings, the estimated impact of the intervention on outcomes in the mathematics achievement domain was neither statistically significant nor large enough to be considered substantively important.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Extent of evidence</th>
<th>Criteria met</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medium to large</td>
<td>Two studies that included 3,062 students across at least 23 schools reported evidence of effectiveness in the mathematics achievement domain.</td>
</tr>
</tbody>
</table>

Two studies that meet WWC group design standards with reservations reported findings in the mathematics achievement domain.

Cai et al. (2011) reported, and the WWC confirmed (after applying a correction for classroom-level clustering), no statistically significant or substantively important difference between the CMP group and the comparison group in the mathematics achievement domain. The WWC characterizes these study findings as an indeterminate effect.

Ridgway et al. (2002) reported a statistically significant difference between the CMP group and the comparison group in the mathematics achievement domain. However, after applying a correction for classroom-level clustering, the WWC found that this difference was no longer statistically significant and the result was not substantively important. The WWC characterizes these study findings as an indeterminate effect.

Thus, for the mathematics achievement domain, neither study showed effects that were statistically significant nor large enough to be considered substantively important. This results in a rating of no discernible effects, with a medium to large extent of evidence.
References

Studies that meet WWC group design standards without reservations

None.

Studies that meet WWC group design standards with reservations


Additional sources:

Additional source:
Studies that do not meet WWC group design standards


Additional sources:


Ellis, J. D. (2011). Middle school mathematics: A study of three programs in south Texas. (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3483008) The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.

Fauth, T. (2007). *Using the Connected Math Project to improve seventh grade math scores at Wapato Middle School* (Unpublished master’s thesis). Heritage University, Toppenish, WA. The study does not meet WWC group design standards because the measures of effectiveness cannot be attributed solely to the intervention.

King, D. A. (2007). *A study to ascertain the effects of the Connected Mathematics Project on student achievement in the Buffalo public schools* (Unpublished master’s thesis). State University of New York at Buffalo. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.


Monaghan, S. R. (2013). Textbooks, teachers, and middle school mathematics student achievement (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 1469609858) The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.


Additional sources:


in Mathematics Education, 39(3), 247–280. The study does not meet WWC group design standards because the measures of effectiveness cannot be attributed solely to the intervention.

**Additional source:**


Winking, D. (1998). The Minneapolis Connected Mathematics Project: Year Two evaluation. Minneapolis, MN: Minneapolis Public Schools. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.

**Additional sources:**


Zvoch, K., & Stevens, J. (2006). Longitudinal effects of school context and practice on middle school mathematics achievement. The Journal of Educational Research, 99(6), 347–357. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.

**Studies that are ineligible for review using the Primary Mathematics Evidence Review Protocol**


Bay, J. M. (1999). Middle school mathematics curriculum implementation: The dynamics of change as teachers introduce and use standards-based curricula. Dissertation Abstracts International, 60(12). This study is ineligible for review because it is out of scope of the protocol.

Bay, J. M., Beem, J. K., Reys, R. E., Papick, I., & Barnes, D. E. (1999). Student reactions to standards-based mathematics curricula: The interplay between curriculum, teachers, and students. School Science and Mathematics, 99(4), 182–188. This study is ineligible for review because it is out of scope of the protocol.

Bay-Williams, J. M., Scott, M. B., & Hancock, M. (2007). Case of the mathematics team: Implementing a team model for simultaneous renewal. The Journal of Educational Research, 100(4), 243–253. This study is ineligible for review because it does not use an eligible design.
Bennett, C. L. (2007). A curriculum project of vocabulary development in the Connected Math program, Moving Straight Ahead (Unpublished master’s thesis). State University of New York College at Brockport. This study is ineligible for review because it is out of scope of the protocol.

Bieda, K. N. (2010). Enacting proof-related tasks in middle school mathematics: Challenges and opportunities. Journal for Research in Mathematics Education, 41(4), 351–382. This study is ineligible for review because it is out of scope of the protocol.

Bledsoe, A. M. (2002). Implementing the Connected Mathematics Project: The interaction between student rational number understanding and classroom mathematical practices. Dissertation Abstracts International, 63(12). This study is ineligible for review because it is out of scope of the protocol.

Booth, J. L., & Koedinger, K. R. (2012). Are diagrams always helpful tools? Developmental and individual differences in the effect of presentation format on student problem solving. British Journal of Educational Psychology, 82(3), 492–511. This study is ineligible for review because it is out of scope of the protocol.

Bray, M. S. (2005). Achievement of eighth grade students in mathematics after completing three years of the Connected Mathematics Project (Unpublished doctoral dissertation). University of Tennessee, Knoxville. This study is ineligible for review because it does not use an eligible design.


Collins, A. M. (2002). What happens to student learning in mathematics when a multi-faceted, long-term professional development model to support standards-based curricula is implemented in an environment of high stakes testing? Dissertation Abstracts International, 65(2). This study is ineligible for review because it does not use an eligible design.

Danielson, C. (2005). Walking a straight line: Introductory discourse on linearity in classrooms and curriculum. Dissertation Abstracts International, 67(2). This study is ineligible for review because it is out of scope of the protocol.

De Groot, C. (2000). Three female voices: The transition to high school mathematics from a reform middle school mathematics program. Dissertation Abstracts International, 61(4). This study is ineligible for review because it is out of scope of the protocol.

Durkin, N. M. (2005). Using Connected Math Program: Its impact on the Delaware State Testing scores of 8th-grade students at Milford Middle School (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 913516241) This study is ineligible for review because it does not use an eligible design.

Ellis, A. B., Özgür, Z., Kulow, T., Williams, C. C., & Amidon, J. (2015). Quantifying exponential growth: Three conceptual shifts in coordinating multiplicative and additive growth. Journal of Mathematical Behavior, 39, 135–155. This study is ineligible for review because it is out of scope of the protocol.

Genz, R. (2006). Determining high school students’ geometric understanding using van Hiele levels: Is there a difference between standards-based curriculum students and non-standards-based curriculum students? (Unpublished master’s thesis). Brigham Young University, Provo, UT. This study is ineligible for review because it is out of scope of the protocol.

Grandau, L., & Stephens, A. C. (2006). Algebraic thinking and geometry. *Mathematics Teaching in the Middle School, 11*(7), 344–349. This study is ineligible for review because it does not use an eligible design.


Griffith, L., Evans, A., & Trowell, J. (2000). *Arkansas grade 8 benchmark exam: How do Connected Mathematics schools compare to state data?* Little Rock: Arkansas State Department of Education. This study is ineligible for review because it is out of scope of the protocol.


Hull, L. S. H. (2000). Teachers’ mathematical understanding of proportionality: Links to curriculum, professional development, and support. *Dissertation Abstracts International, 62*(2). This study is ineligible for review because it does not use an eligible design.

Izsak, A. (2008). Mathematical knowledge for teaching fraction multiplication. *Cognition and Instruction, 26*(1), 95–143. This study is ineligible for review because it does not use an eligible design.


Izsak, A., Tillema, E., & Tunç-Pekkan, Z. (2008). Teaching and learning fraction addition on number lines. *Journal for Research in Mathematics Education, 39*(1), 33–62. This study is ineligible for review because it is out of scope of the protocol.


Keiser, J. M. (1997). The development of students’ understanding of angle in a non-directive learning environment. *Dissertation Abstracts International, 58*(8). This study is ineligible for review because it is out of scope of the protocol.

Keiser, J. M. (2010). Shifting our computational focus. *Mathematics Teaching in the Middle School, 16*(4), 216–223. This study is ineligible for review because it does not use an eligible design.

Krebs, A. S. (2003). Middle grades students’ algebraic understanding in a reform curriculum. *School Science and Mathematics, 103*(5), 233–245. This study is ineligible for review because it does not use an eligible design.

**Additional source:**


Kulm, G., Morris, K., & Grier, L. (1999). *Middle grades mathematics textbooks: A benchmarks-based evaluation.* Washington, DC: American Association for the Advancement of Science. This study is ineligible for review because it is out of scope of the protocol.


Newton, J. A. (2012). Investigating the mathematical equivalence of written and enacted middle school standards-based curricula: Focus on rational numbers. *International Journal of Educational Research, 51*, 66–85. This study is ineligible for review because it does not use an eligible design.


Prentice Hall. (2006). CMP: Research and evaluation summary. Upper Saddle River, NJ: Author. This study is ineligible for review because it does not use an eligible design.


Quigley, D. (2010). *Project-based learning and student achievement* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 741546820) This study is ineligible for review because it does not use an eligible design.


Sinclair, N., & Armstrong, A. (2011). Tell a piecewise story. *Mathematics Teaching in the Middle School, 16*(6), 346–353. This study is ineligible for review because it does not use an eligible design.


Stevens, B. B. A. (2005). *The development of pedagogical content knowledge of a mathematics teaching intern: The role of collaboration, curriculum, and classroom context* (Unpublished doctoral dissertation). University of Missouri–Columbia. This study is ineligible for review because it is out of scope of the protocol.


Appendix A.1: Research details for Cai et al. (2011)


Additional sources:


Table A1. Summary of findings

<table>
<thead>
<tr>
<th>Outcome domain</th>
<th>Sample size</th>
<th>Average improvement index (percentile points)</th>
<th>Statistically significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematics achievement</td>
<td>14 schools/606 students</td>
<td>+4</td>
<td>No</td>
</tr>
</tbody>
</table>

Setting

The study was conducted in 14 middle schools, all of which were located in a single large, urban school district in the United States.

Study sample

Seven middle schools that implemented CMP in grades 6–8 were selected for the study. Another seven middle schools in the district that were not implementing CMP were selected for the study's comparison group, based on their similarity to the CMP schools on demographic characteristics. The study sample consisted of students who began sixth grade in fall 2005. Students in both the CMP and comparison groups were assessed in fall 2005 (at the beginning of sixth grade), and in spring 2006, spring 2007, and spring 2008 (at the end of each grade). The eighth-grade analytic sample, assessed in spring 2008, consisted of 606 students with an equal number of students and schools in the intervention and comparison groups. The analytic sample was about half the size of the baseline sample, which included 1,284 students. About 85% of the students in the baseline sample were minorities: 64% were African American, 16% were Hispanic, 4% were Asian, and 1% were Native American. The remaining 15% of the students were White.

In addition, the CMP and comparison students were tracked into high school and outcomes were assessed at the end of each grade in ninth, tenth, and eleventh grades (spring 2009, spring 2010, and spring 2011). In the 10 high schools that were included in this follow-up sample, the CMP and comparison students were mixed together in the same classrooms and used the same (non-CMP) curricula.

The eighth-grade findings are considered the main outcomes in this review and presented in Appendix C because they are the most immediate outcome measuring the 3 full years of CMP use. The sixth-, seventh-, and eleventh-grade outcomes are considered supplemental findings presented in Appendix D that do not factor into the intervention's rating of effectiveness.11
**Intervention group**

Students in the intervention schools used the first edition of CMP (version 1) as their core mathematics curriculum in the sixth, seventh, and eighth grades in the 2005–06 through 2007–08 school years; specific details about how CMP was implemented in study schools are not provided by the authors.

**Comparison group**

Students in the comparison schools used one of several traditional mathematics curricula already in use in their schools during each grade (grades 6, 7, and 8); specific details about how the comparison curricula were implemented are not provided by the authors. The authors conducted detailed analyses on the curricular materials to examine differences between CMP and one of the curricula used by the comparison group (Glencoe). The authors noted that there were differences between CMP and non-CMP curriculum; notably, CMP emphasizes problem solving while the non-CMP curricula take a more traditional approach that focuses on concepts and procedures. The authors also indicate that there were some differences between the different non-CMP curricula, but the differences between the comparison curricula were not substantial since they took the same traditional approach to math instruction. The authors did not name the other math curricula used by students in the comparison group.

**Outcomes and measurement**

The study included two outcome measures that meet WWC review requirements and fall within the mathematics achievement domain: (a) open-ended tasks and (b) ability to pose problems. Both assessments were developed by the researchers involved in the study. The first assessment was administered four times: at the beginning of sixth grade (fall 2005), at the end of sixth grade (spring 2006), at the end of seventh grade (spring 2007), and at the end of eighth grade (spring 2008). The second outcome was measured at the end of eleventh grade (spring 2011). For a more detailed description of the outcome measures, see Appendix B. The study also examined five outcomes that do not meet WWC standards. Three middle school assessments were used in analyses that did not meet standards because the CMP and comparison groups were not found to be equivalent at baseline: (a) translation, (b) computation, and (c) equation solving. Two additional outcomes, the Classroom Assessment Based on Standards (CABS) administered in ninth grade and the state mathematics achievement test administered in tenth grade, do not meet standards because the authors did not provide evidence of baseline equivalence for the analytic sample.

Five other outcomes used by the authors were not eligible for review. Three middle school outcomes are ineligible because they are measures of implementation fidelity: (a) the level of conceptual and procedural emphasis in lessons, (b) the difficulty of instructional tasks, and (c) the difficulty of homework problems. Two eleventh-grade outcomes are ineligible because they were used in an analysis that drew on an ineligible design (posttest only): a graphing task and equation solving task.

**Support for implementation**

The authors did not provide any information on support for implementation.
Appendix A.2: Research details for Ridgway et al. (2002)


Table A2. Summary of findings

<table>
<thead>
<tr>
<th>Outcome domain</th>
<th>Sample size</th>
<th>Average improvement index (percentile points)</th>
<th>Study findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematics achievement</td>
<td>9 sites/2,456 students</td>
<td>0</td>
<td>No</td>
</tr>
</tbody>
</table>

Setting

The study was conducted in nine sites across the United States (five in the Midwest, two in the West, and two in the East). The authors do not indicate whether a site is a single school or school district.

Study sample

The study sample consisted of sixth- and seventh-grade students in the 1994–95 school year and eighth-grade students in the 1995–96 school year. The intervention and comparison group participants were matched to the extent possible on ability, location, and diversity in student population. In five of the nine sites that participated in the study, only a small number of teachers were using CMP, so comparison classrooms were selected locally. At the four other sites, comparison classrooms were identified in alternate locations. At each site, pairs of classrooms were selected within each grade level to form the intervention group; one comparison classroom was selected for every pair of intervention classrooms. The 1994–95 sample included 338 sixth-grade students and 627 seventh-grade students from 36 classrooms (18 in each grade) who used the CMP curriculum and 234 seventh-grade students from 18 comparison group classrooms (nine in each grade). The 1995–96 sample included 820 eighth-grade students from 14 classrooms using CMP and 275 students from seven comparison classrooms. The authors provided results by grade.

For this review, student data were combined across grades; the effectiveness rating is based on the combined analyses for each outcome measure and presented in Appendix C. Although some intervention students in this combined analysis used CMP in a previous school year, the combined finding measures the effectiveness of receiving 1 year of the intervention because the pre-intervention measures were assessed at the beginning of same school year in which outcomes were measured. The authors did not report demographic characteristics of the study students.

Intervention group

Students in the intervention group used CMP as their core math curriculum. Specific details about how CMP was implemented in study schools are not provided by the authors. The sixth- and seventh-grade intervention students used CMP in the 1994–95 school year, and the eighth-grade intervention students used CMP in the 1995–96 school year. The sixth-grade students had no prior use of CMP; however, approximately three-fourths of the seventh- and eighth-grade students had used CMP in the previous year. The authors did not indicate which edition of CMP was used, but it was likely the first edition of CMP, since the study was conducted between 1994–96 and the second edition of CMP was not developed until 2000.
**Comparison group**

Students in the comparison group used commercially available mathematics textbooks. The authors did not provide the name of the comparison texts, nor did they provide details about how the comparison curricula were implemented in study schools. Teachers in the comparison group did not use the 

**Outcomes and measurement**

The study included two outcome measures, both of which meet the review requirements and fall within the math achievement domain: (a) the Iowa Test of Basic Skills (ITBS, a standardized test) and (b) the Balanced Assessment (BA) test (developed as a collaboration between the study authors and the Balanced Assessment Project). Both tests were administered to the study students in the fall and spring of the 1994–95 school year for sixth- and seventh-grade students and in the fall and spring of the 1995–96 school year for eighth-grade students.

The grade-level analyses for the ITBS and eighth-grade analyses for the BA do not demonstrate equivalence of the analytic intervention and comparison groups. Only the sixth- and seventh-grade-level analyses for the BA test demonstrate equivalence. These BA outcomes are presented as supplemental findings in Appendix D. For a more detailed description of both outcome measures, see Appendix B.

**Support for implementation**

All CMP teachers attended a summer CMP workshop at Michigan State University. This workshop included sessions that involved teachers experiencing the curriculum as students as well as sessions to share methods and techniques for implementation. The authors indicate that they do not have information on how CMP materials were used in the classroom.
### Appendix B: Outcome measures for the mathematics achievement domain

<table>
<thead>
<tr>
<th>Mathematics achievement</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Balanced Assessment (BA) test</strong></td>
<td>The BA test was developed in collaboration between the Balanced Assessment Project and study authors to assess reasoning, mathematical communication, mathematical problem solving, and the ability to make connections among mathematical concepts. The BA test does not mimic the language or content of the CMP curriculum and is therefore not considered overaligned with the intervention. The BA instrument assessed a variety of curricular topics and consisted of constructed-response items that required a range of responses: items required students to interpret real-world situations, explain their strategies, and provide justifications for their responses. Each item varied in the number of points awarded for a correct answer. Five different forms of the BA test were used. BA tests were scored by trained study staff (which included mathematics teachers, mathematics education graduate students, senior high school students, and professionals who attended a structured training) using rubrics and guidelines under the direction of the CMP and Balanced Assessment Project staff. The BA tests had a reported inter-rater reliability of at least 90% (as cited in Ridgway et al., 2002). This outcome is only reported as a supplemental finding in Appendix D.</td>
</tr>
<tr>
<td><strong>Iowa Test of Basic Skills (ITBS)</strong></td>
<td>The ITBS includes five subtests (Concepts, Estimation, Problem Solving, Data Interpretation, and Computation) which collectively assess numbers and operations skills and concepts, as in other standardized tests. The test includes 60 timed multiple-choice items, each worth one point. The ITBS tests were scored by the Riverside Testing Company (as cited in Ridgway et al., 2002).</td>
</tr>
<tr>
<td><strong>Open-ended tasks</strong></td>
<td>This test included students' scores on five open-ended tasks (combined into one score) adapted from the BA. The tasks tested students' high-level thinking skills as well as procedural knowledge and reasoning in problem-solving. Two middle school math teachers were trained to code and score the tests. The exact agreement between two coders was approximately 80% and agreement within one point was over 95%. The authors reported internal consistency (Cronbach's Alpha) for each measurement occasion, ranging from .65 to .77 across forms and measurement occasions (as cited in Cai et al., 2011).</td>
</tr>
<tr>
<td><strong>Problem-posing ability</strong></td>
<td>Problem-posing ability was a researcher-developed task that involved identifying an equation for a given graph and posing a real-life situation that could be represented by the graph. Students' answers were scored by two researchers using a rubric. Agreement between coders ranged from 92% to 100% (as cited in Cai et al., 2011). This outcome is only reported as a supplemental finding in Appendix D.</td>
</tr>
</tbody>
</table>
### Appendix C: Findings included in the rating for the mathematics achievement domain

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Study sample</th>
<th>Sample size</th>
<th>Mean (standard deviation)</th>
<th>WWC calculations</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Intervention group</td>
<td>Comparison group</td>
<td>Mean difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cai et al. (2011)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Open-ended tasks total score</em></td>
<td>Grade 8</td>
<td>14 schools/606 students</td>
<td>575.00 (95.00)</td>
<td>565.00 (100.00)</td>
<td>10.0</td>
</tr>
<tr>
<td><strong>Domain average for mathematics achievement</strong> (Cai et al., 2011)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.10</td>
</tr>
<tr>
<td><strong>Ridgway et al. (2002)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Balanced Assessment (BA)</em></td>
<td>Grades 6, 7</td>
<td>9 sites/1,361 students</td>
<td>nr</td>
<td>nr</td>
<td>na</td>
</tr>
<tr>
<td><em>Iowa Test of Basic Skills (ITBS)</em></td>
<td>Grades 6, 7, 8</td>
<td>9 sites/2,456 students</td>
<td>8.78 (2.88)</td>
<td>8.77 (2.97)</td>
<td>0.01</td>
</tr>
<tr>
<td><strong>Domain average for mathematics achievement</strong> (Ridgway et al., 2002)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Domain average for mathematics achievement across all studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.05</td>
</tr>
</tbody>
</table>

**Table Notes:** For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of the study's domain average was determined by the WWC. Some statistics may not sum as expected due to rounding. na = not available. nr = not reported.

* For Cai et al. (2011), means and standard deviations in the table were obtained through an author query. A correction for clustering was needed but did not affect whether any of the contrasts were found to be statistically significant. The p-value presented here was reported in the original study. The WWC calculated the intervention group mean using a difference-in-differences approach by adding the impact of the intervention (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. This study is characterized as having an indeterminate effect because the estimated effect is neither statistically significant nor substantively important. For more information, please refer to the WWC Standards and Procedures Handbook (version 3.0), p. 26.

* For Ridgway et al. (2002), the authors reported unadjusted ITBS means and standard deviations separately by grades 6, 7, and 8; the grades were pooled together by the WWC. The BA did not demonstrate equivalence when pooled across grades 6, 7, and 8, nor did it demonstrate equivalence for the grade 8 subgroup; therefore, the WWC pooled together grades 6 and 7. For the BA, the authors reported p-values using the results from an ANCOVA model, but did not report information needed to calculate a WWC effect size. The p-values presented here were reported in the original study. Corrections for clustering and multiple comparisons were needed, but the authors did not provide enough information to determine WWC significance for the BA; the WWC-computed p-value is .97 for the ITBS; therefore, the WWC does not find the result to be statistically significant. The WWC calculated the ITBS intervention group mean using a difference-in-differences approach by adding the impact of the intervention (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. The WWC excludes findings without an effect size from the domain averages. This study is characterized as having an indeterminate effect because the estimated effect is neither statistically significant nor substantively important. For more information, please refer to the WWC Standards and Procedures Handbook (version 3.0), p. 26.
# Appendix D: Description of supplemental findings for the mathematics achievement domain

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Study sample</th>
<th>Sample size</th>
<th>Mean (standard deviation)</th>
<th>WWC calculations</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Intervention group</td>
<td>Comparison group</td>
<td>Mean difference</td>
</tr>
<tr>
<td>Cai et al. (2011)&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open-ended tasks total score</td>
<td>Grade 6</td>
<td>14 schools/606 students</td>
<td>494.00 (97.00)</td>
<td>502.00 (97.00)</td>
<td>−8.00</td>
</tr>
<tr>
<td>Open-ended tasks total score</td>
<td>Grade 7</td>
<td>14 schools/606 students</td>
<td>538.00 (92.00)</td>
<td>531.00 (93.00)</td>
<td>7.00</td>
</tr>
<tr>
<td>Problem-posing performance</td>
<td>Grade 11 (3-year follow-up)</td>
<td>136 students</td>
<td>1.56 (1.14)</td>
<td>0.50 (1.17)</td>
<td>1.06</td>
</tr>
<tr>
<td>Problem-posing performance</td>
<td>Grade 11, middle third subgroup (3-year follow-up)</td>
<td>45 students</td>
<td>0.27 (0.87)</td>
<td>0.05 (0.23)</td>
<td>0.22</td>
</tr>
<tr>
<td>Problem-posing performance</td>
<td>Grade 11, bottom third subgroup (3-year follow-up)</td>
<td>45 students</td>
<td>0.11 (0.72)</td>
<td>0.00 (0.00)</td>
<td>0.11</td>
</tr>
</tbody>
</table>

| Ridgway et al. (2002)<sup>b</sup> |              |             |                           |                  |          |          |               |         |
| Balanced Assessment test | Grade 6 | 27 classrooms/500 students | nr | nr | na | na | na | <.001 |
| Balanced Assessment test | Grade 7 | 27 classrooms/861 students | nr | nr | na | na | na | <.001 |

**Table Notes:** The supplemental findings presented in this table are additional findings from studies in this report that meet WWC design standards with or without reservations, but do not factor into the determination of the intervention rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual’s percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding. na = not available. nr = not reported.

<sup>a</sup> For Cai et al. (2011), the unadjusted means and standard deviations reported in this table were obtained through an author query. The authors did not report p-values with these data. Corrections for multiple comparisons were not implemented because the authors did not report p-values and the WWC-computed p-values did not indicate any of these findings were statistically significant. The eighth-grade findings are considered the main outcomes in this review and are presented in Appendix C because they are the most immediate outcomes measuring the 3 full years of CMP use. The sixth- and seventh-grade findings (which represent 1 and 2 years of CMP use, respectively) and eleventh-grade outcomes (which represent a 3 year follow-up) that are presented in this table are considered supplemental findings. For the eleventh-grade outcomes, the author provided data separately for the bottom, middle, and top thirds of students. The WWC examined results separately by the achievement subgroups, and combined data to estimate the effect for the full eleventh-grade sample. The author examined different eleventh-grade outcomes using two different baseline tests: an open-ended task and an equation-solving task. The full eleventh-grade sample demonstrated baseline equivalence on the open-ended task but not on the equation-solving task. The subgroup analyses for the eleventh grade only met WWC baseline equivalence standards on the open-ended task for the middle third achievement level subgroup and the equation-solving test for the bottom third achievement subgroup. The WWC calculated the intervention group mean using a difference-in-differences approach by adding the impact of the intervention (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means for all outcomes except the eleventh-grade problem-posing performance (bottom third subgroup). Please see the WWC Procedures and Standards Handbook (version 3.0) for more information.

<sup>b</sup> For Ridgway et al. (2002), the authors reported p-values using the results from an ANCOVA model, but did not report adjusted means and standard deviations. Corrections for clustering and multiple comparisons were needed but did not affect whether any of the contrasts were found to be statistically significant. The p-values presented here were reported in the original study.
Endnotes

1 The descriptive information for this intervention was obtained from a publicly available source: the intervention’s website (https://connectedmath.msu.edu). The WWC requests developers review the intervention description sections for accuracy from their perspective. The intervention description was provided to the developer in June 2015; however, the WWC received no response. Further verification of the accuracy of the descriptive information for this intervention is beyond the scope of this review.

2 The WWC previously released a report on CMP under the Middle School Mathematics topic area in January 2010; the report was prepared using the WWC Procedures and Standards Handbook (version 1.0) and the Middle School Mathematics review protocol (version 1.0). In June 2015, the WWC restructured the reviews of research on math interventions into two areas instead of three. These two review areas are Primary Mathematics (which includes interventions in which math is presented through multi-topic materials and curricula, typically used in grades K–8), and Secondary Mathematics (which includes interventions that are organized by math content area [e.g., Algebra, Geometry, and Calculus], typically taught in grades 9–12). These two areas replaced the prior Elementary School Math, Middle School Math, and High School Math areas, which were organized by student grade level. The WWC is updating and replacing intervention reports written under the prior topic areas.

The literature search for the current report reflects documents publicly available by March 2016. This report has been updated to include reviews of 40 studies that were not included in the prior report. Of the additional studies, 31 were not within the scope of the review protocol for the Primary Mathematics topic area, and eight were within the scope of the review protocol for the Primary Mathematics topic area but did not meet WWC group design standards. One study meets WWC group design standards with reservations, and findings from this study are summarized in this report. A complete list and disposition of all studies reviewed are provided in the references.

The report includes reviews of all studies included in the previous report and resulted in a revised disposition for two studies. Ridgway et al. (2002) received a disposition in this report of meets WWC group design standards with reservations, where it had previously received the rating of does not meet WWC group design standards. The study received the previous rating because the analytic sample included some students that had prior use of the intervention. This review is based on the Primary Mathematics topic area review protocol (version 3.1), which does not prohibit prior use of an intervention, as long as the study demonstrates baseline equivalence on the analytic sample and meets all other review requirements. Using the WWC Procedures and Standards Handbook (version 3.0) and the Primary Mathematics review protocol (version 3.1), several analytic contrasts within this study meet WWC group design standards, and therefore, the study is now rated meets WWC group design standards with reservations.

Schneider (2000) received a disposition in this report of does not meet WWC group design standards, where it had previously received a rating of meets WWC group design standards with reservations. This study has baseline differences in the adjustment range (0.05 to 0.25 standard deviations), and therefore must include an appropriate statistical adjustment. However, an appropriate statistical adjustment was not conducted by the author. Therefore, the prior rating was incorrect and is corrected in this report. The study is now rated does not meet WWC group design standards.

3 The studies in this report were reviewed using the standards from the WWC Procedures and Standards Handbook (version 3.0), and the Primary Mathematics review protocol (version 3.1). The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

4 This represents the total analytic sample size across the two studies. Cai et al. (2011) included students from 14 schools within one school district. Ridgway et al. (2002) included students from nine sites across the country; the authors did not define whether a site is a single school, a school district, or a larger entity.

5 Please see the Primary Mathematics review protocol (version 3.1) for more information about the outcome domain.

6 For criteria used in the determination of the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on p. 23. These improvement index numbers show the average and range of individual-level improvement indices for all findings across the studies.

7 This review does not include any studies of CMP Algebra I since it is not eligible for review under the WWC’s Primary Mathematics review protocol.

8 One study that meets standards and is summarized in this report provides evidence of effectiveness about the first edition of CMP. The other study that meets standards and is summarized in this report may also provide evidence of effectiveness about the first edition of CMP. While the author does not specify the edition used, due to the timing of the study, it likely provides evidence about the first edition of CMP.

9 For the analyses of eleventh-grade outcomes, the authors followed students who enrolled in the 10 high schools with the largest number of CMP and non-CMP students from the original sample. The eleventh-grade outcomes included 135–136 students from the original sample of nearly 700 students. The number of classes from which the eleventh-grade sample originated was not reported by
the author. In high school, the CMP and non-CMP students were no longer separated and they used the same curriculum (which was not CMP).

10 The WWC identified three additional sources related to Cai et al. (2011). These studies do not contribute unique information to Appendix A.1 and are not listed here.

11 The ninth- and tenth-grade analytic samples did not demonstrate equivalence as required, and therefore do not meet WWC group design standards; thus, they are not presented in Appendix D.

12 The WWC identified one additional source related to Ridgway et al. (2002). The study does not contribute unique information to Appendix A.2 and is not listed here.

13 Although the study presented findings for each grade separately, the WWC combined this data when presenting the findings in Appendix C. Within each study and outcome domain, the WWC reports a set of primary findings from multiple outcomes that use comparable samples, when possible. When the WWC combined the data across all three grades for each outcome measure, baseline equivalence was demonstrated on the ITBS, but not on the BA. Therefore, we report as the primary findings the combined grades 6 and 7 sample for the BA, and the combined sample across all three grades for the ITBS. In addition, the two grade subgroups that demonstrated equivalence on the BA are presented as supplementary results in Appendix D.

Recommended Citation

### WWC Rating Criteria

**Criteria used to determine the rating of a study**

<table>
<thead>
<tr>
<th>Study rating</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meets WWC group design standards without reservations</td>
<td>A study that provides strong evidence for an intervention's effectiveness, such as a well-implemented RCT.</td>
</tr>
<tr>
<td>Meets WWC group design standards with reservations</td>
<td>A study that provides weaker evidence for an intervention's effectiveness, such as a QED or an RCT with high attrition that has established equivalence of the analytic samples.</td>
</tr>
</tbody>
</table>

**Criteria used to determine the rating of effectiveness for an intervention**

<table>
<thead>
<tr>
<th>Rating of effectiveness</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive effects</td>
<td>Two or more studies show statistically significant positive effects, at least one of which met WWC group design standards for a strong design, AND No studies show statistically significant or substantively important negative effects.</td>
</tr>
<tr>
<td>Potentially positive effects</td>
<td>At least one study shows a statistically significant or substantively important positive effect, AND No studies show a statistically significant or substantively important negative effect AND fewer or the same number of studies show indeterminate effects than show statistically significant or substantively important positive effects.</td>
</tr>
<tr>
<td>Mixed effects</td>
<td>At least one study shows a statistically significant or substantively important positive effect AND at least one study shows a statistically significant or substantively important negative effect, but no more such studies than the number showing a statistically significant or substantively important positive effect, OR At least one study shows a statistically significant or substantively important effect AND more studies show an indeterminate effect than show a statistically significant or substantively important effect.</td>
</tr>
<tr>
<td>Potentially negative effects</td>
<td>One study shows a statistically significant or substantively important negative effect and no studies show a statistically significant or substantively important positive effect, OR Two or more studies show statistically significant or substantively important negative effects, at least one study shows a statistically significant or substantively important positive effect, and more studies show statistically significant or substantively important negative effects than show statistically significant or substantively important positive effects.</td>
</tr>
<tr>
<td>Negative effects</td>
<td>Two or more studies show statistically significant negative effects, at least one of which met WWC group design standards for a strong design, AND No studies show statistically significant or substantively important positive effects.</td>
</tr>
<tr>
<td>No discernible effects</td>
<td>None of the studies shows a statistically significant or substantively important effect, either positive or negative.</td>
</tr>
</tbody>
</table>

**Criteria used to determine the extent of evidence for an intervention**

<table>
<thead>
<tr>
<th>Extent of evidence</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medium to large</td>
<td>The domain includes more than one study, AND The domain includes more than one school, AND The domain findings are based on a total sample size of at least 350 students, OR, assuming 25 students in a class, a total of at least 14 classrooms across studies.</td>
</tr>
<tr>
<td>Small</td>
<td>The domain includes only one study, OR The domain includes only one school, OR The domain findings are based on a total sample size of fewer than 350 students, AND, assuming 25 students in a class, a total of fewer than 14 classrooms across studies.</td>
</tr>
</tbody>
</table>
Glossary of Terms

**Attrition**
Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.

**Clustering adjustment**
If intervention assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.

**Confounding factor**
A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.

**Design**
The design of a study is the method by which intervention and comparison groups were assigned.

**Domain**
A domain is a group of closely related outcomes.

**Effect size**
The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.

**Eligibility**
A study is eligible for review and inclusion in this report if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.

**Equivalence**
A demonstration that the analysis sample groups are similar on observed characteristics defined in the review area protocol.

**Extent of evidence**
An indication of how much evidence supports the findings. The criteria for the extent of evidence levels are given in the WWC Rating Criteria on p. 23.

**Improvement index**
Along a percentile distribution of individuals, the improvement index represents the gain or loss of the average individual due to the intervention. As the average individual starts at the 50th percentile, the measure ranges from –50 to +50.

**Intervention**
An educational program, product, practice, or policy aimed at improving student outcomes.

**Intervention report**
A summary of the findings of the highest-quality research on a given program, product, practice, or policy in education. The WWC searches for all research studies on an intervention, reviews each against design standards, and summarizes the findings of those that meet WWC design standards.

**Multiple comparison adjustment**
When a study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.

**Quasi-experimental design (QED)**
A quasi-experimental design (QED) is a research design in which study participants are assigned to intervention and comparison groups through a process that is not random.

**Randomized controlled trial (RCT)**
A randomized controlled trial (RCT) is an experiment in which eligible study participants are randomly assigned to intervention and comparison groups.

**Rating of effectiveness**
The WWC rates the effects of an intervention in each domain based on the quality of the research design and the magnitude, statistical significance, and consistency in findings. The criteria for the ratings of effectiveness are given in the WWC Rating Criteria on p. 23.

**Single-case design**
A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.
Glossary of Terms

Standard deviation  The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample tend to be spread out over a large range of values.

Statistical significance  Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < .05$).

Substantively important  A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.

Systematic review  A review of existing literature on a topic that is identified and reviewed using explicit methods. A WWC systematic review has five steps: 1) developing a review protocol; 2) searching the literature; 3) reviewing studies, including screening studies for eligibility, reviewing the methodological quality of each study, and reporting on high quality studies and their findings; 4) combining findings within and across studies; and, 5) summarizing the review.

Please see the WWC Procedures and Standards Handbook (version 3.0) for additional details.
An intervention report summarizes the findings of high-quality research on a given program, practice, or policy in education. The WWC searches for all research studies on an intervention, reviews each against evidence standards, and summarizes the findings of those that meet standards.

This intervention report was prepared for the WWC by Mathematica Policy Research under contract ED-IES-13-C-0010.