

What Works Clearinghouse



University of Chicago School Mathematics Project 6–12 Curriculum

Program Description¹

The *University of Chicago School Mathematics Project (UCSMP) 6–12 Curriculum* is a series of yearlong courses—(1) Transition Mathematics; (2) Algebra; (3) Geometry; (4) Advanced Algebra; (5) Functions, Statistics, and Trigonometry; and (6) Precalculus and Discrete Mathematics—emphasizing problem solving, real-world applications, and the use of technology. The program

is designed to allow schools to offer the appropriate math to students regardless of grade level. Beginning with the Algebra course, technology is used in the classroom to aid in the development of properties and skills, and graphing calculators are used to complete assignments at all levels.

Research²

Two studies of *UCSMP* that fall within the scope of the High School Math review protocol meet What Works Clearinghouse (WWC) evidence standards with reservations. Together, the two studies included 251 high school students in five schools in five districts.³

Based on these two studies, the WWC considers the extent of evidence for *UCSMP* on high school students to be small for math achievement because the ratings are based on reviews with fewer than 350 students in total.

Effectiveness

UCSMP was found to have potentially positive effects on math achievement for high school students.

	Math achievement
Rating of effectiveness	Potentially positive
Improvement index⁴	Average: +23 percentile points Range: +2 to +38 percentile points

1. The descriptive information for this program was obtained from publicly available sources: the developer’s website (<http://ucsmc.uchicago.edu>, downloaded April 2010) and the publisher’s website (<http://www.mheonline.com>, downloaded August 2010). The WWC requests developers to review the program description sections for accuracy from their perspective. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review. The literature search reflects documents publicly available by October 2010.
2. The studies in this report were reviewed using WWC Evidence Standards, Version 2.0 (see the WWC Procedures and Standards Handbook, Chapter III), as described in protocol Version 2.0.
3. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.
4. These numbers show the average and range of student-level improvement indices for all findings across the studies.

Additional program information

Developer and contact

Developed by the University of Chicago School Mathematics Project (UCSMP), the third edition of the *University of Chicago School Mathematics Project 6–12 Curriculum* is distributed by Wright Group/McGraw-Hill. Email: SEG_customerservice@mcgraw-hill.com. Web: <http://www.mheonline.com>. Telephone: (800) 523-2371.

For additional information on the program, contact either Zalman Usiskin, professor emeritus of education and director, UCSMP (Address: 6030 South Ellis Avenue, Chicago, IL 60637. Email: z-usiskin@uchicago.edu. Telephone: (773) 702-1560), or Denisse Thompson, professor of mathematics education, University of South Florida (Address: 4202 East Fowler Ave., #DU105, Tampa, FL 33620-5650. Email: denisse@usf.edu. Telephone: (813) 974-2687).

Scope of use

The first edition of *UCSMP* was developed and tested beginning in 1985 and the second edition beginning in 1992.⁵ The third edition was developed and tested between 2005 and 2007 and is available through the Wright Group/McGraw-Hill. According to the developers at the University of Chicago School Mathematics Project, more than 3.5 million students in elementary, middle, and high schools are using *UCSMP* materials and curricula.

Research

Twenty studies reviewed by the WWC High School Math topic area investigated the effects of any of the six topic areas of *UCSMP*. Two studies (Hirschhorn, 1993; Thompson, Senk, Witonsky, Usiskin, & Kaeley, 2006) are quasi-experimental designs that meet WWC evidence standards with reservations. The remaining 18 studies do not meet either WWC evidence standards or eligibility screens.

5. The two studies meeting standards with reservations in this report focus on different editions of the program. Per Thompson et al. (2006), the major differences between the first and second editions are an increased use of technology and emphasis on student writing and projects.

Teaching

Each course of the *UCSMP* includes the following components: the student textbook (hardcover), teacher's edition (including the electronic teacher's edition, eTe), teacher resources, assessment resources, and technology resources. Lessons in the student book contain activities, full examples, and partially completed guided examples to model skills and problem solving. Students are encouraged to assess their own understanding with an End-of-Chapter Self-Test correlated to objectives. Projects provided at the end of each chapter are designed as extended activities, giving students experience using real data. The use of technology—including graphing calculators at all grade levels, geometry systems, spreadsheets, the Internet, and other computer applications—is considered an essential component of the curricula.

Cost

For each individual curriculum, the student textbook costs \$63.00. A bundled, complete teacher resource package consisting of the teacher's edition and teacher resources (volumes 1 and 2), assessment CD-ROM, and electronic teacher's edition with answers and solutions (volumes 1 and 2) costs \$346.50. See the publisher's website for pricing of individual resource items.

Meets evidence standards with reservations

Hirschhorn (1993) conducted a longitudinal, four-year quasi-experimental evaluation of the *UCSMP* (first edition) in three high schools in which both traditional and *UCSMP* curricula were used. Every student in the intervention group received four years of *UCSMP* curricula, starting with Transition Mathematics in 7th grade and ending with *UCSMP* Advanced Algebra in 10th grade. No student in the comparison group participated in any *UCSMP* courses, instead enrolling in math classes using other curricula.

Research *(continued)*

Treatment students were matched to potential comparison students retrospectively using pretest (6th-grade) scores, and outcomes were measured at the end of the four-year period (10th grade). Two high schools (B and C in the study), located in affluent suburbs with predominantly White, higher-achieving students, contributed a total of 62 students to the analysis. A third high school (A in the study) is excluded from the WWC review because baseline equivalence between the intervention and comparison groups could not be established on math scores alone.

In the second study meeting evidence standards with reservations, Thompson et al. (2006) used students' math ability to match pairs of classrooms within a school to evaluate the effectiveness of the second edition of *UCSMP Algebra*. Random assignment of teachers, within each pair, to the treatment or control condition was not always possible; therefore, this study is treated as a quasi-experimental design. The WWC review focuses on the three

high schools (X, Y, and Z in the study) that compared the use of the *UCSMP* second edition to that of non-*UCSMP* textbooks.⁶ These three high schools, located in three different regions of the country and serving three different populations of students, contributed 189 primarily 9th-grade students to the analysis.

Extent of evidence

The WWC categorizes the extent of evidence in each domain as small or medium to large (see the WWC Procedures and Standards Handbook, Appendix G). The extent of evidence takes into account the number of studies and the total sample size across the studies that meet WWC evidence standards with or without reservations.⁷

The WWC considers the extent of evidence for *UCSMP* to be small for math achievement for high school students because the ratings are based on reviews with fewer than 350 students in total (see Appendix A5).

Effectiveness Findings

The WWC review of interventions for High School Math addresses student outcomes in one domain: math achievement. The findings below present the authors' estimates and WWC-calculated estimates of the size and the statistical significance of the effects of *UCSMP* on high school students.⁸

Math achievement. Two studies presented findings in the math achievement domain. Using the Mathematics Level 1 Achievement test, Hirschhorn (1993) reports no statistically significant effect in one school, and a positive and statistically significant effect in another. WWC calculations pooled samples

across the two sites and found differences that were neither statistically significant (at the 0.05 level) nor large enough to be considered substantively important according to WWC criteria (i.e., an effect size of at least 0.25). Based on a second measure, the *UCSMP*-developed Applications Test, Hirschhorn (1993) reports, and the WWC confirms, positive and statistically significant effects. Thompson et al. (2006) report no statistically significant effects on the standardized High School Subjects Test: Algebra, but positive and statistically significant effects on both the *UCSMP*-developed Algebra⁹ and Problem-Solving tests. The WWC confirmed the statistical significance of the latter findings,

6. Another analysis included schools in which the comparison group used *UCSMP Algebra*, first edition. This design is not eligible for review based on the topic area review protocol.
7. The extent of evidence categorization was developed to tell readers how much evidence was used to determine the intervention rating, focusing on the number and size of studies. Additional factors associated with a related concept—external validity, such as the students' demographics and the types of settings in which studies took place—are not taken into account for the categorization. Information about how the extent of evidence rating was determined for *UCSMP* is in Appendix A5.
8. The level of statistical significance was reported by the study authors or, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. In the case of Hirschhorn (1993), a correction for multiple comparisons was needed, so the significance levels may differ from those reported in the original study. In the case of Thompson et al. (2006), corrections for clustering and multiple comparisons were needed, so the significance levels may differ from those reported in the original study.
9. Care must be taken in interpreting results for the *UCSMP*-developed Algebra test, as comparison students in school Y had an opportunity to learn the content for less than one-quarter of the test items.

Effectiveness (*continued*) using clustering and multiple comparison corrections. In sum, both studies demonstrated statistically significant positive effects in the math achievement domain, using small samples of students and only on developer-created instruments.¹⁰

Rating of effectiveness

The WWC rates the effects of an intervention in a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative. The rating of effectiveness takes into account four factors: the quality of the research design, the statistical significance of the findings, the size of the difference between participants in the intervention and the comparison conditions, and the consistency in findings across studies (see the WWC Procedures and Standards Handbook, Appendix E).

The WWC found UCSMP to have potentially positive effects on math achievement for high school students.

Improvement index

The WWC computes an improvement index for each individual finding. In addition, within each outcome domain, the WWC computes an average improvement index for each study and an

average improvement index across studies (see WWC Procedures and Standards Handbook, Appendix F). The improvement index represents the difference between the percentile rank of the average student in the intervention condition and the percentile rank of the average student in the comparison condition. Unlike the rating of effectiveness, the improvement index is entirely based on the size of the effect, regardless of the statistical significance of the effect, the study design, or the analysis. The improvement index can take on values between -50 and $+50$, with positive numbers denoting favorable results for the intervention group.

The average improvement index for math achievement is $+23$ percentile points across the two studies, with a range of $+2$ to $+38$ percentile points across findings.

Summary

The WWC High School Math topic area reviewed 20 studies on the *UCSMP*. Two of these studies meet WWC evidence standards with reservations; the remaining 18 studies do not meet either WWC evidence standards or eligibility screens. Based on the two studies, the WWC found potentially positive effects in math achievement for high school students. The conclusions presented in this report may change as new research emerges.

References

Meets WWC evidence standards with reservations

Hirschhorn, D. B. (1993). A longitudinal study of students completing four years of UCSMP mathematics. *Journal for Research in Mathematics Education*, 24(2), 136–158.

Additional source:

Hirschhorn, D. B. (1992). *Implementation of first four years of the UCSMP secondary curriculum* (Unpublished doctoral dissertation). University of Chicago, IL.

Thompson, D. R., Senk, S. L., Witonsky, D., Usiskin, Z., & Kaeley, G. (2006). *An evaluation of the second edition of UCSMP Algebra*. Chicago, IL: University of Chicago School Mathematics Project.

Studies that fall outside the High School Math review protocol or do not meet WWC evidence standards

Davis, J. D., & Shih, J. C. (2007). Secondary options and post-secondary expectations: Standards-based mathematics programs and student achievement on college mathematics placement exams. *School Science and Mathematics*, 107(8), 336–346. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Hedges, L. V., Stodolsky, S., Flores, P. V., Mathison, D., Mathison, S., Sarther, C. S., & Zhang, J. (1988). Formative evaluation of USCMP Advanced Algebra. *Advanced Algebra*

10. These instruments demonstrated validity and reliability as discussed in Appendix A2.

References (continued)

- teacher's edition (pp. T46–T48). Chicago, IL: University of Chicago School Mathematics Project. The study is ineligible for review because it does not occur within the time frame specified in the protocol.
- Hedges, L. V., Stodolsky, S., Mathison, S., & Flores, P. V. (1986). *Transition Mathematics field study*. Chicago, IL: University of Chicago School Mathematics Project. Summarized in *Transition Mathematics teacher's edition* (1990, pp. T46–T50). Chicago, IL: University of Chicago School Mathematics Project. The study is ineligible for review because it does not occur within the time frame specified in the protocol.
- Henderson, B. K. (1996). An evaluation of the use of UCSMP materials in a mathematics program (University of Chicago School Math Project). *Dissertation Abstracts International*, 57(11A), 188-4637. (AAG9714356). The study is ineligible for review because it does not include an outcome within a domain specified in the protocol.
- Hirschhorn, D. B., & Senk, S. (1992). Calculators in the UCSMP curriculum for grades 7 and 8. In J. T. Fey & C. R. Hirsch (Eds.), *Calculators in mathematics education*. Reston, VA: National Council of Teachers of Mathematics. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Mathison, S., Hedges, L. V., Stodolsky, S., Flores, P., & Sarther, C. (1989). *Teaching and learning algebra: An evaluation of UCSMP Algebra* (Evaluation Rep. No. 88/89-ALG-1). Chicago, IL: University of Chicago. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.
- McConnell, J. (1990). *UCSMP sophomores on the PSAT*. Glenview, IL: Glenbrook South High School. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.
- Senk, S. L. (1989). Assessing students' knowledge of functions. In C. A. Mahrer, G. A. Goldin, & R. B. Davis (Eds.), *Proceedings of the eleventh annual meeting of the North American chapter of the International Group for the Psychology of Mathematics Education* (vol. 2). New Brunswick, NJ: Rutgers University. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.
- Senk, S. L. (2003). Effects of the UCSMP secondary curriculum on students' achievement. In S. L. Senk & D. R. Thompson (Eds.), *Standards-based school mathematics curricula: What are they? What do students learn?* (pp. 425–456). Mahwah, NJ: Lawrence Erlbaum. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Senk, S. L., & Thompson, D. R. (2006). Brief report: Strategies used by second-year algebra students to solve problems. *Journal for Research in Mathematics Education*, 37, 116–128. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.
- Swann, J. M. (1996). *An investigation into the effectiveness of transition mathematics*. Paper presented at the American Educational Research Association Annual Conference, New York, NY. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Thompson, D. R. (1994). *An evaluation of a new course in pre-calculus and discrete mathematics* (Unpublished doctoral dissertation). University of Chicago, IL. The study does not meet WWC evidence standards because it includes only outcomes that are overaligned with the intervention or measured in a way that is inconsistent with the protocol.

References *(continued)*

- Thompson, D. R., & Senk, S. L. (1993). Assessing reasoning and proof in high school. In N. L. Webb & A. Coxford (Eds.), *1993 yearbook of the National Council of Teachers of Mathematics* (pp. 167–176). Reston, VA: National Council of Teachers of Mathematics. The study is ineligible for review because it does not examine the effectiveness of an intervention.
- Thompson, D. R., & Senk, S. L. (2001). The effects of curriculum on achievement in second-year algebra: The example of the University of Chicago School Mathematics Project. *Journal for Research in Mathematics Education*, 32(1), 58. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.
- Thompson, D. R., Senk, S. L., Witonsky, D., Usiskin, Z., & Kaeley, G. (2001). *An evaluation of the second edition of UCSMP Advanced Algebra*. Chicago, IL: University of Chicago School Mathematics Project. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.
- Thompson, D. R., Senk, S. L., Witonsky, D., Usiskin, Z., & Kaeley, G. (2005). *An evaluation of the second edition of UCSMP Transition Mathematics*. Chicago, IL: University of Chicago School Mathematics Project. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Thompson, D. R., Witonsky, D., Senk, S. L., Usiskin, Z., & Kaeley, G. (2003). *An evaluation of the second edition of UCSMP Geometry*. Chicago, IL: University of Chicago School Mathematics Project. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.
- Zahrt, L. T. (2001). *High school reform math programs: An evaluation for leaders* (Doctoral dissertation). Eastern Michigan University, Ypsilanti. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.

Appendix

Appendix A1.1 Study characteristics: Hirschhorn, 1993

Characteristic	Description
Study citation	Hirschhorn, D. B. (1993). A longitudinal study of students completing four years of UCSMP mathematics. <i>Journal for Research in Mathematics Education</i> , 24(2), 136–158.
Participants	<p>After identifying a set of locations that used four years of <i>UCSMP</i> curricula, the researcher focused on three places where both traditional and <i>UCSMP</i> curricula were used. Every student in the intervention group received four years of <i>UCSMP</i> curricula, starting with Transition Mathematics in 7th grade and ending with <i>UCSMP</i> Advanced Algebra. No student in the control group received any <i>UCSMP</i> courses.</p> <p>It is uncertain how students were initially assigned to the different courses, so the researcher matched students on a pretreatment outcome ex post facto. Since the intervention began in 7th grade, students were matched on their 6th-grade performance on a standardized exam (in 1986).¹ The process differed slightly by site. In site A, the only available data on 6th-grade academic performance were composite scores for math, reading, and general logic; because baseline equivalence on math achievement alone could not be established, this site is excluded from WWC analysis. In sites B and C, students were essentially matched on 6th-grade math scores and then on 6th-grade reading scores. Since some of the control students in sites B and C had not taken advanced algebra by the end of 10th grade, two comparison groups were formed. We focused on the age-based comparison sample, wherein the comparison group was based on all students who started 7th grade at the same time (and thus were in 10th grade at the time of the posttest, in 1990).² The eligible sample for this intervention report was 62 students (31 intervention and 31 comparison): 26 students in site B (13 intervention and 13 comparison) and 36 students in site C (18 intervention and 18 comparison).³</p> <p>We pooled results across sites B and C because when sites were analyzed separately, all students in the treatment group in site B had the same teacher and class in at least one grade, thus generating a confound.</p>
Setting	This study took place in three locations that were unspecified due to confidentiality concerns. Sites B and C were affluent suburbs with predominantly White, relatively high-performing students (in the 80th to 90th percentile nationally in both reading and math). In site B, students were selected from a single class in the largest of seven feeder schools serving the district high school. In site C, students were selected from four classes in three feeder schools to the high school serving the district.
Intervention	The intervention students used <i>UCSMP</i> math curricula for four consecutive years, starting in 7th grade and ending in 10th grade. They used Transition Mathematics in 1986–87, <i>UCSMP</i> Algebra in 1987–88, <i>UCSMP</i> Geometry in 1988–89, and <i>UCSMP</i> Advanced Algebra in 1989–90.
Comparison	The comparison students used traditional math curricula that correspond to each of the <i>UCSMP</i> courses, produced by Addison-Wesley, Merrill, Houghton Mifflin, and others.
Primary outcomes and measurement	The Mathematics Level I Achievement Test, Form 3JAC2 (College Board, 1988), a standardized eligible outcome, was measured in 10th grade, near the end of the final academic year (April/May 1990). The developer-created math achievement test, administered during the same period, also was considered. For a more detailed description of these outcome measures, see Appendix A2.
Staff/teacher training	There was no specific staff or teacher training specified. However, at site B, some of the 7th- and 8th-grade teachers were involved with the <i>UCSMP</i> pilot, so they had previous experience working with the curricula.

1. There are potential concerns with matching a comparison sample based on 6th-grade test scores when the posttest is taken in 10th grade. This is especially problematic when the treatment group consists of students who have persisted in the treatment for four consecutive years, which is likely to be a select sample of initial entrants (when even the initial entrants could be a self-selected group), compared with a retrospectively chosen comparison group.
2. The second course-based comparison group used students at the end of advanced algebra. Since that could involve comparing 10th-grade treatment students with 11th-grade comparison students, we did not include this comparison in the intervention report.
3. We excluded the 22 students in site A (11 intervention students and 11 comparison students).

Appendix A1.2 Study characteristics: Thompson et al., 2006

Characteristic	Description
Study citation	Thompson, D. R., Senk, S. L., Witonsky, D., Usiskin, Z., & Kaeley, G. (2006). <i>An evaluation of the second edition of UCSMP Algebra</i> . Chicago, IL: University of Chicago School Mathematics Project.
Participants	<p>Schools were recruited for the study by advertising in UCSMP and National Council of Teachers of Mathematics publications. To participate, a school needed at least four sections of the equivalent of an algebra class, at either middle or high school, and had to promise to keep classes intact for a full year. The selected schools provided the names of at least two teachers willing to participate in the study. The study utilized a matched-pairs design, in which classes were matched in the same school on the basis of students' math ability. When possible, teachers were randomly assigned to treatment conditions within each pair; however, local conditions did not always permit random assignment.</p> <p>The treatment group used <i>UCSMP Algebra</i> second edition, whereas the comparison group used either <i>UCSMP Algebra</i> first edition or a non-<i>UCSMP</i> comparison textbook. The former comparison was not eligible for review based on the topic area review protocol. This review focuses on the three high schools—X, Y, and Z as labeled by the authors—that compared the <i>UCSMP</i> second edition to non-<i>UCSMP</i> textbooks. Analysis sample size was 189 students (98 intervention and 91 comparison): 28 in school X (14 intervention and 14 comparison), 114 in school Y (65 intervention and 49 comparison), and 47 in school Z (19 intervention and 28 comparison).¹ The majority of these students (n=161) were in 9th grade; the rest were enrolled in grades 10–12.</p>
Setting	The three high schools included in the results were located on the West Coast, in the Northeast, and in the South. School X was a large, ethnically diverse high school on the West Coast, serving roughly 2,800 students in grades 9–12 from inner-city and suburban environments; the only <i>UCSMP</i> text previously used at the school was <i>Geometry</i> . School Y was a suburban high school in the Northeast, serving 950 students in grades 9–12 from a middle- to upper-middle-class socioeconomic population; no previous <i>UCSMP</i> texts were used at this school. School Z was a suburban high school of about 2,800 students in grades 9–12, which was located in a middle- to upper-middle-class neighborhood in the South and serves a large Hispanic community; no previous <i>UCSMP</i> texts were used in this school.
Intervention	Teachers assigned to the treatment condition used <i>UCSMP Algebra</i> (second edition, field trial version) during the 1992–93 school year.
Comparison	The non- <i>UCSMP</i> classes used the following algebra texts at the time of the study: <i>Algebra I: An Incremental Development</i> (Saxon) in School X, Houghton Mifflin's <i>Algebra: Structure and Method Book I</i> in School Y, and Prentice Hall's <i>Algebra I</i> in School Z.
Primary outcomes and measurement	Shortly before the end of the school year, the teachers administered several instruments: (1) High School Subjects Tests: Algebra, (2) a developer-created Algebra Test, and (3) a developer-created Algebra Problem-Solving and Understanding Test. The last test was administered in two forms, with half the students in each class randomly assigned to the even form of the test and the other half assigned to the odd form. For a more detailed description of these outcome measures, see Appendix A2.
Staff/teacher training	No direct in-service was provided to the teachers using <i>UCSMP Algebra</i> either before or during the school year. Although teachers had a tentative table of contents for the entire book when the school year began, they received the actual text in three spiral-bound parts: chapters 1–4 at the beginning of the school year, chapters 5–8 around November, and chapters 9–13 in early winter. Additionally, teachers received lesson notes and answers to questions, one chapter at a time, throughout the school year. To assist with the formative evaluation, the teachers completed a chapter evaluation form after completing each chapter. These teachers also met in Chicago once in the fall and again in the spring to give feedback to the developers about the materials; during these meetings, there were brief opportunities to raise issues related to the use of technology and the use of reading and group problem solving in class and to discuss other instructional concerns.

1. We excluded 334 students from the analysis comparing the *UCSMP Algebra* second edition to the first edition.

Appendix A2 Outcome measures for the math achievement domain

Outcome measure	Description
Mathematics Level I Achievement Test, Form 3JAC2 (College Board, 1988)	This 50-question, multiple-choice, standardized achievement instrument published by the College Board covers geometry and second-year algebra content (as cited in Hirschhorn, 1993). The test was graded on the College Board range of 200 to 800 points. On the reliability tests for the samples in the study, the coefficient alpha ranged from 0.601 to 0.925 (median of 0.873).
UCSMP Applications Test	This 30-question, multiple-choice test was developed by UCSMP for this study as well as for use by future researchers and school personnel (as cited in Hirschhorn, 1993). Scientific calculators were provided and their use was encouraged. Items were selected from the <i>Second International Mathematics Study</i> (Chang & Ruzicka, 1985), the <i>Formative Study of UCSMP Advanced Algebra</i> (Hedges et al., 1988), a submittal to the College Board by personal letter in 1988, and some original items. On the reliability tests for the samples in the study, the coefficient alpha ranged from 0.637 to 0.842 (median of 0.739).
High School Subjects Test: Algebra (American Testronics, 1988)	This 40-question, multiple-choice standardized test (as cited in Thompson et al., 2006) focuses on algebraic concepts. Calculators are not allowed during the test. For samples in the study, the Kuder-Richardson KR20 was approximately 0.80.
UCSMP Algebra Test	This 40-item, multiple-choice test was developed by UCSMP to assess algebraic content in <i>UCSMP Algebra</i> as well as topics important in all algebra classes regardless of curriculum (as cited in Thompson et al., 2006). Calculators were permitted during the test. For samples in the study, the KR20 was between 0.81 and 0.83.
UCSMP Problem-Solving and Understanding Test	This open-ended problem-solving test was administered in two forms, with half the students in each class randomly assigned to the even form of the test and the other half assigned to the odd form (as cited in Thompson et al., 2006). Each form contained four open-ended items, and item-specific rubrics were developed for each item. Each item was scored independently and blindly by two raters. Inter-rater reliabilities were as follows: Even Form: 78.8% (item 1), 89.7% (item 2), 74.2% (item 3), 93.1% (item 4a), 89.4% (item 4b) Odd Form: 79.3% (item 1), 82.9% (item 2), 88.1% (item 3), 92.8% (item 4a), 94.8% (item 4b)

Appendix A3 Summary of study findings included in the rating for the math achievement domain¹

Outcome measure	Study sample	Sample size (classes/ students)	Authors' findings from the study		WWC calculations			
			Mean outcome (standard deviation) ²		Mean difference ⁵ (UCSMP – comparison)	Effect size ⁶	Statistical significance ⁷ (at $\alpha = 0.05$)	Improvement index ⁸
			UCSMP group ³	Comparison group ⁴				
Hirschhorn, 1993⁹								
Mathematics Level I Achievement	Grade 10, sites B & C	62 students	492.26 (46.39)	489.03 (78.71)	3.23	0.05	ns	+2
UCSMP Applications	Grade 10, sites B & C	62 students	21.10 (3.88)	16.46 (3.95)	4.64	1.17	Statistically significant	+38
Average for math achievement (Hirschhorn, 1993)¹⁰						0.61	Statistically significant	+23
Thompson et al., 2006⁹								
High School Subjects: Algebra	Mostly grade 9, schools X, Y, & Z	12 classes/ 189 students	47.9 (16.3)	46.0 (14.9)	1.9	0.12	ns	+5
UCSMP Algebra	Mostly grade 9, schools X, Y, & Z	12 classes/ 189 students	49.5 (16.3)	37.3 (14.9)	12.2	0.78	Statistically significant	+28
UCSMP Problem-Solving and Understanding ¹¹	Mostly grade 9, schools X, Y, & Z	12 classes/ 189 students	6.23 (3.69)	3.39 (2.54)	2.84	0.89	Statistically significant	+31
Average for math achievement (Thompson et al., 2006)¹⁰						0.59	Statistically significant	+22
Domain average for math achievement across all studies¹⁰						0.60	na	+23

ns = not statistically significant

na = not applicable

(continued)

Appendix A3 Summary of study findings included in the rating for the math achievement domain¹ (continued)

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the math achievement domain.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. The intervention group values are pooled across the eligible sites.
4. The comparison group values are pooled across the eligible sites.
5. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
6. For an explanation of the effect size calculation, see WWC Procedures and Standards Handbook, Appendix B.
7. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
8. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between -50 and +50, with positive numbers denoting favorable results for the intervention group.
9. The level of statistical significance was reported by the study authors or, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. In the case of Hirschhorn (1993), a correction for multiple comparisons was needed, so the significance levels may differ from those reported in the original study. In the case of Thompson et al. (2006), corrections for clustering and multiple comparisons were needed, so the significance levels may differ from those reported in the original study.
10. The WWC-computed average effect sizes for each study and for the domain across studies are simple averages rounded to two decimal places. The average improvement indices are calculated from the average effect sizes.
11. The group values for each condition are pooled across even and odd test forms.

Appendix A4 UCSMP rating for the math achievement domain

The WWC rates an intervention's effects in a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.¹

For the outcome domain of math achievement, the WWC rated *UCSMP* as potentially positive for high school students. The remaining ratings (mixed, no discernible effects, potentially negative, and negative) were not considered, as *UCSMP* was assigned the highest applicable rating.

Rating received

Potentially positive effects: Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

Met. Two studies of *UCSMP* showed a statistically significant positive effect.

AND

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

Met. No studies of *UCSMP* showed indeterminate or statistically significant negative effects.

Other ratings considered

Positive effects: Strong evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *positive* effects, at least one of which met WWC evidence standards for a *strong* design.

Not met. Although two studies of *UCSMP* showed statistically significant positive effects, neither met WWC evidence standards for a strong design.

AND

- Criterion 2: No studies showing statistically significant or substantively important *negative* effects.

Met. No studies of *UCSMP* showed statistically significant negative effects.

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain-level effect. The WWC also considers the size of the domain-level effect for ratings of potentially positive or potentially negative effects. For a complete description, see the WWC Procedures and Standards Handbook, Appendix E.

Appendix A5 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence ¹
		Schools	Students	
Math achievement	2	5	251	Small

1. A rating of “medium to large” requires at least two studies and two schools across studies in one domain, and a total sample size across studies of at least 350 students or 14 classrooms. Otherwise, the rating is “small.” For more details on the extent of evidence categorization, see the WWC Procedures and Standards Handbook, Appendix G.