

# Appendix

## Appendix A1 Study characteristics: Schneider, 2000 (quasi-experimental design)

Characteristic	Description
<b>Study citation</b>	Schneider, C. L. (2000). <i>Connected Mathematics and the Texas Assessment of Academic Skills</i> (Doctoral dissertation, University of Texas at Austin, 2000). <i>Dissertation Abstracts International</i> , 61(12), 4709A. (UMI No. 9997642)
<b>Participants</b>	The study included three cohorts from 23 <i>CMP</i> schools and 25 matched comparison schools overall. However, because baseline equivalence was established only for cohorts 1 and 2 through 1998, the WWC excluded cohort 3 from this review. <sup>1</sup> Cohort 1 included more than 3,000 <i>CMP</i> students and 2,600 comparison students. Cohort 2 included more than 3,400 <i>CMP</i> students and 2,900 comparison students.
<b>Setting</b>	The participating schools were located in rural, suburban, and urban, as well as both low and high socioeconomic, areas of Texas. The schools varied in the English language learner status of the student populations that they served. Many of the schools had predominantly minority student populations.
<b>Intervention</b>	Schools in the treatment group used <i>CMP</i> , starting with grade 6 in 1996–97, adding grade 7 in 1997–98, and adding grade 8 in 1998–99. Cohorts in the school-level analysis represented all students in grades using <i>CMP</i> , regardless of whether the students were enrolled every year of implementation. By 1998, students in cohort 1 had received the intervention for up to two years (grades 6 and 7), and students in cohort 2 had received the intervention for one year (grade 6). There was substantial variation in the extent to which the curriculum was used at each grade level and each year across these schools.
<b>Comparison</b>	Schools in the comparison group did not use <i>CMP</i> . The author did not report the mathematics curricula used by comparison schools. The 25 comparison schools were matched to treatment schools using a regression analysis of variables that predicted 1996 Texas Assessment of Academic Skills (TAAS) passing rates.
<b>Primary outcomes and measurement</b>	The primary outcome measure included in this review was the school-level passing rate on the mathematics portion of the TAAS. For a more detailed description of this outcome measure, see Appendix A2.
<b>Staff/teacher training</b>	In the summer prior to implementation, teachers in the <i>CMP</i> schools participated in a six-day summer professional development provided by the Texas Statewide System Initiative. The training discussed units and emphasized mathematical understanding and pedagogy. Many teachers also participated in a two-day follow-up professional development in the summer after implementation.

1. The author conducted both student-level and school-level analyses. However, baseline equivalence was established only for the school-level analysis. In addition, the sample of students for whom achievement scores were reported changed over time—special education students were included in 1999 achievement data but were excluded in prior years. Because baseline equivalence was established only for regular education students, the WWC excluded the 1999 data from this review. This exclusion of the 1999 data also resulted in the exclusion of cohort 3 because 1999 is the only year for which posttest data were reported for cohort 3.

## Appendix A2 Outcome measure for the math achievement domain

Outcome measure	Description
<b>Texas Assessment of Academic Skills (TAAS) pass rate</b>	The TAAS is a criterion-referenced test that measures problem-solving and critical-thinking skills. The measure used is the percentage of students that passed the math portion of the TAAS (as cited in Schneider, 2000).

## Appendix A3 Summary of study findings included in the rating for the math achievement domain<sup>1</sup>

Outcome measure	Study sample	Sample size (schools/students)	Author's findings from the study		WWC calculations			
			Mean outcome (standard deviation) <sup>2</sup>		Mean difference <sup>3</sup> (CMP – comparison)	Effect size <sup>4</sup>	Statistical significance <sup>5</sup> (at $\alpha = 0.05$ )	Improvement index <sup>6</sup>
			CMP group	Comparison group				
<b>Schneider, 2000 (quasi-experimental design)<sup>7</sup></b>								
TAAS pass rate	Grades 6 & 7 (Cohorts 1 & 2) <sup>7</sup>	96/>12,162	82% <sup>8</sup> of students (na)	82% <sup>9</sup> of students (na)	0.00	0.00	ns	0
<b>Average for math achievement (Schneider, 2000)<sup>9</sup></b>						<b>0.00<sup>10</sup></b>	<b>ns</b>	<b>0</b>
<b>Domain average for math achievement across all studies<sup>9</sup></b>						<b>0.00</b>	<b>na</b>	<b>0</b>

ns = not statistically significant

na = not applicable

TAAS = Texas Assessment of Academic Skills

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the math achievement domain. Separate cohort findings from Schneider (2000) are not included in these ratings, but are reported in Appendix A4.
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. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
4. For an explanation of the effect size calculation, see WWC Procedures and Standards Handbook, Appendix B.
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. 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 results favorable to the intervention group.
7. The WWC calculated a sample-weighted average of the cohort 1 and 2 effects to use in the intervention rating. Cohort 1, the 7th grade cohort, had the intervention for two years and cohort 2, the 6th grade cohort, had the intervention for one year. Separate findings for cohorts 1 and 2 are reported in Appendix A4.
8. The intervention group value from Schneider (2000) is the unadjusted comparison group mean plus the difference in mean gains between the intervention (CMP) and comparison groups.
9. The comparison group mean from Schneider (2000) is unadjusted.
10. The WWC used school-level pass rates to calculate the effect size because student-level pass rates were not available. In general, the school-level figures should serve as a reasonable approximation of the individual figures if there are not large differences in pass rates between schools of substantially different enrollments. Based on prior pass rates (1996) provided by the author in the appendix, the WWC concluded that the pass rates computed using school-level data and student-level data were not substantially different. As a result, based on the available school-level pass rates, the WWC estimated that the mean difference of 0 percentage points resulted in an effect size of 0.

## Appendix A4 Summary of cohort findings for the math achievement domain<sup>1</sup>

Outcome measure	Study sample	Sample size (schools/students)	Author's findings from the study					
			Mean outcome (standard deviation) <sup>2</sup>		WWC calculations			
			CMP group <sup>3</sup>	Comparison group <sup>4</sup>	Mean difference <sup>5</sup> (CMP – comparison)	Effect size <sup>6</sup>	Statistical significance <sup>7</sup> (at $\alpha = 0.05$ )	Improvement index <sup>8</sup>
<b>Schneider, 2000 (quasi-experimental design)<sup>9</sup></b>								
TAAS pass rate	Grade 7 (Cohort 1)	48/>5,701	0.80 (na)	0.80 (na)	0.00 (na)	0.00	ns	0
TAAS pass rate	Grade 6 (Cohort 2)	48/>6,461	0.84 (na)	0.83 (na)	0.01 (na)	0.06	ns	+2

ns = not statistically significant

na = not applicable

TAAS = Texas Assessment of Academic Skills

1. This appendix presents separate cohort findings for measures that fall in math achievement. A sample-weighted average of the cohort effects was used for rating purposes and is presented in Appendix A3.
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 value from Schneider (2000) is the unadjusted comparison group mean plus the difference in mean gains between the intervention (CMP) and comparison groups.
4. The comparison group mean from Schneider (2000) is unadjusted.
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 results favorable to 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 an explanation about the clustering correction, see the WWC Tutorial on Mismatch. For the formulas the WWC used to calculate 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 Schneider (2000), no corrections for clustering or multiple comparisons were needed.

## Appendix A5 CMP rating for the math achievement domain

The WWC rates an intervention's effects for a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.<sup>1</sup> For the outcome domain of math achievement, the WWC rated *CMP* as having no discernible effects.

### Rating received

**No discernible effects:** No affirmative evidence of effects.

- Criterion 1: None of the studies show a statistically significant or substantively important effect, either *positive* or *negative*.

**Met.** None of the studies showed statistically significant or substantively important positive or 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.** No studies showed statistically significant positive effects.

### AND

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

**Met.** No studies showed statistically significant or substantively important negative effects.

**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.

**Not met.** No studies showed statistically significant or substantively important positive effects.

### 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.

**Not met.** The one study that evaluated math achievement and met WWC standards showed indeterminate effects.

**Mixed effects:** Evidence of inconsistent effects as demonstrated through either of the following criteria.

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

**Not met.** No studies showed statistically significant or substantively important effects, either positive or negative.

### OR

- Criterion 2: At least one study showing a statistically significant or substantively important effect, and more studies showing an *indeterminate* effect than showing a statistically significant or substantively important effect.

**Not met.** No studies showed statistically significant or substantively important effects, either positive or negative.

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**     **CMP rating for the math achievement domain** *(continued)*

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

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

**Not met.** No studies showed statistically significant or substantively important negative effects.

**AND**

- Criterion 2: No studies showing a statistically significant or substantively important *positive* effect, or more studies showing statistically significant or substantively important *negative* effects than showing statistically significant or substantively important *positive* effects.

**Met.** No studies showed statistically significant or substantively important positive effects.

**Negative effects:** Strong evidence of a negative effect with no overriding contrary evidence.

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

**Not met.** No studies showed statistically significant negative effects.

**AND**

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

**Met.** No studies showed statistically significant or substantively important positive effects.

## Appendix A6 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence <sup>1</sup>
		Schools	Students	
Math achievement	1	96	>12,162 <sup>2</sup>	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.
2. Schneider (2000) reported the number of students only for the student-level analysis. The student-level analysis included only students who were at the same school every year of implementation, whereas the school-level analysis reviewed in this report includes all students in the schools, regardless of whether they were enrolled every year. The WWC used the number of students in the student-level analysis as the minimum for the number of students in the school-level analysis: 5,701 students in cohort 1 and 6,461 students in cohort 2.