

Appendix

Appendix A1.1 Study Characteristics: Nunnery & Ross, 2007 (quasi-experimental design)

Characteristic	Description
Study citation	Nunnery, J. A., & Ross, S. M. (2007). The effects of the School Renaissance program on student achievement in reading and mathematics. <i>Research in the Schools</i> , 14(1), 40–59.
Participants	The analysis sample included 992 students (482 treatment, 510 comparison) in the 2001/02 grade 8 cohort from four middle schools (two treatment and two comparison). Of the student sample, 21% qualified for free or reduced-price lunch (21% treatment, 20% comparison), 4% were limited English proficient (4% treatment, 4% comparison), 7% African-American (9% treatment, 6% comparison), 3% Asian (2% treatment, 3% comparison), 19% Hispanic (19% treatment, 20% comparison), 0% Native American (0% treatment, 0% comparison), and 70% were White (69% treatment, 71% comparison). Information about attrition was provided only at the level of assignment. Of the 11 elementary and middle schools originally selected as comparison schools, three schools did not provide appropriate grade-level test score data and were replaced (it is unknown whether any of these replaced schools were middle schools). Students in the analysis sample remained in the same school and had matched data available for three consecutive years (1999/2000–2001/02).
Setting	The treatment group schools came from one suburban school district in Texas. Comparison schools came from other school districts in Texas with similar populations of students.
Intervention	In 2000/01, schools in the treatment group implemented School Renaissance, a comprehensive school reform model that includes <i>Accelerated Math</i> . <i>Accelerated Math</i> is a progress-monitoring software program that tracks students' daily activities, provides immediate feedback to students and teachers, alerts teachers to students struggling with certain assignments, and monitors achievement. Teachers can use the program with their existing math curriculum. Students in the treatment group experienced two years of the <i>Accelerated Math</i> program.
Comparison	Schools in the comparison condition were from Texas school districts that had not implemented the full School Renaissance package. However, it is still possible that some elements of <i>Accelerated Math</i> were present in the comparison schools.
Primary outcomes and measurement	The study used the Texas Learning Index (TLI) math scores (based on the Texas Assessment of Academic Skills); for the grade 8 cohort, program comparisons were based on average transformed scores for grades 7 and 8 from 2001 and 2002. The TLI has a common interpretation across grades: a score of 70 or above indicates performance at or above grade-level expectations. A student receiving the same score at consecutive grade levels made one year of academic progress. For a more detailed description of these outcome measures, see Appendix A2.
Staff/teacher training	A Renaissance coach conducts an initial training seminar and provides ongoing assistance to teachers.

Appendix A1.2 Study Characteristics: Ysseldyke & Bolt, 2007 (randomized controlled trial with attrition)

Characteristic	Description
Study citation	Ysseldyke, J., & Bolt, D. M. (2007). Effect of technology-enhanced continuous progress monitoring on math achievement. <i>School Psychology Review</i> , 36(3), 453–467.
Participants¹	The initial study sample included 3,309 students in grades 2–8 during the 2003/04 school year from 133 classrooms in nine schools, representing eight school districts in eight states. In the initial study sample 1% of the students were Asian, 28% African-American, 38% Hispanic, 0% Native American, 24% White, and 8% not specified. This review focuses on the middle school sample, which initially included 1,823 grade 6–8 students (1,010 treatment and 813 control) in 73 classrooms (41 treatment and 32 control). Demographic data on the middle school students could not be culled from the original study. Middle school classrooms dropped from the analysis include: 7 special education or enrichment treatment classrooms taught by teachers who had access to, but did not receive training in, <i>Accelerated Math</i> ; 4 classrooms (2 treatment, 2 control) taught by two teachers who, according to the authors, arbitrarily chose which students to treat; and 22 classrooms (11 treatment, 11 controls) in a large, urban middle school district that, according to the authors, was unable to devote sufficient time and resources to <i>Accelerated Math</i> . The results here are drawn from the test-takers in the 40 middle school classrooms (21 treatment, 19 control) included in the analysis—792 students took the STAR Math test (418 treatment, 374 control) and 851 took the Terra Nova test (454 treatment, 397 control). Postattrition treatment and control groups were equivalent on pretests at baseline. Because these samples reflect attrition rates greater than 20%, the WWC rated this study as meeting evidence standards with reservations.
Setting	The study took place in eight schools in seven districts in seven states: Alabama, Florida, Michigan, Mississippi, North Carolina, South Carolina, and two schools in Texas. The middle school sample analyzed here comprises three schools in Michigan, Mississippi, and North Carolina.
Intervention	Students were taught by teachers using the <i>Accelerated Math</i> program during the 2003/04 school year. <i>Accelerated Math</i> is a progress-monitoring software program that teachers can use with their existing math curriculum. The program tracks students' daily activities, provides immediate feedback to students and teachers, alerts teachers to students struggling with certain assignments, and monitors student achievement. Teachers assigned to the treatment group were asked to use <i>Accelerated Math</i> with their present math curriculum. In practice, the program was not implemented for approximately 40% of grade 2–8 students in the initial treatment group; the authors did not report the percentage of grade 6–8 students in the treatment group of the analysis sample that did not participate in <i>Accelerated Math</i> .
Comparison	Students in the control group were taught using existing math curricula, without <i>Accelerated Math</i> . The existing curricula included: Scott Foresman Middle School Math, Consumer Math, Everyday Math, Transition Math (Prentice Hall), and Chicago Math in Michigan; Glencoe in Mississippi; and Glencoe, McGraw-Hill, and the state curriculum in North Carolina. Control students had the same teachers as the intervention group students.
Primary outcomes and measurement	Participating students were pretested in October 2003 and posttested in May 2004 using two nationally normed, standardized tests (STAR Math and Terra Nova) for math achievement. Students in the treatment and control groups were compared using a linear regression analysis in which posttest scores were regressed on pretest scores and on dummy variables related to main effects for experimental condition and school.
Staff/teacher training	Teachers in the intervention group were trained to use <i>Accelerated Math</i> . During the school year, teachers using <i>Accelerated Math</i> received three to five visits from a Renaissance Learning math consultant, who guided teachers on how to improve their use of the program. Teachers also had unlimited access to technical support.

1. The study authors provided the WWC with sample sizes for the middle schools.

Appendix A1.3 Study Characteristics: Ysseldyke & Tardrew, 2007 (quasi-experimental design)

Characteristic	Description
Study citation	Ysseldyke, J., & Tardrew, S. (2007). Use of a progress monitoring system to enable teachers to differentiate mathematics instruction. <i>Journal of Applied School Psychology, 24</i> (1), 1–28.
Participants¹	The initial study sample included 2,397 students (1,319 treatment and 1,078 comparison) in grades 3–10 during the 2001/02 school year from 125 classrooms (67 treatment and 58 comparison) in 47 schools in 24 states. The middle school analysis sample in this review included 475 grade 6–8 students (235 treatment, 240 comparison) in 25 classrooms (13 treatment, 12 comparison). Of the students, 43% were male (43% treatment, 43% comparison), and 49% female (48% treatment, 51% comparison). Of the total student gender, 7% were reported as unspecified (8% treatment, 6% comparison). Of the students, 0% were Asian (0% treatment, 0% comparison), 1% African-American (1% treatment, 0% comparison), 9% Hispanic (9% treatment, 9% comparison), 0% Native American (0% treatment, 0% comparison), 44% White (38% treatment, 49% comparison), and 46% were reported as unspecified (51% treatment, 42% comparison).
Setting	The study was conducted in 47 schools in 24 states (Alabama, Arkansas, California, Georgia, Idaho, Illinois, Indiana, Iowa, Kansas, Massachusetts, Michigan, Minnesota, Missouri, Montana, New Mexico, Ohio, Oklahoma, Oregon, Pennsylvania, Tennessee, Texas, Virginia, Washington, and Wisconsin). The authors did not report whether all schools and states were represented in the middle school (grades 6–8) sample.
Intervention	Students were taught by teachers using <i>Accelerated Math</i> during the spring semester of the 2001/02 school year. A progress-monitoring software program, <i>Accelerated Math</i> can be used with teachers' existing math curriculum. The program tracks students' daily activities, provides immediate feedback to students and teachers, alerts teachers' to students struggling with certain assignments, and monitors student achievement. Teachers assigned to the <i>Accelerated Math</i> treatment group were asked to use the program with their existing math curriculum.
Comparison	Teachers assigned to the comparison group did not use <i>Accelerated Math</i> but continued their usual math curriculum and practices.
Primary outcomes and measurement	Using a computer adaptive test of math achievement (STAR Math), students were pretested in January 2002 and posttested in May 2002.
Staff/teacher training	Intervention teachers participated in a one-day training session conducted by Renaissance Learning. The training was designed to familiarize teachers with <i>Accelerated Math</i> and to guide them in integrating it into curriculum and instruction. Of 68 treatment group teachers in the full sample, 66 attended the training.

1. The study authors provided the WWC with the number of middle school classrooms by treatment status.

Appendix A2 Outcome measures for the math achievement domain

Outcome measure	Description
Texas Learning Index math scores (based on the Texas Assessment of Academic Skills)	The Texas Assessment of Academic Skills (TAAS) is a criterion-referenced standardized state test that measures problem-solving and critical-thinking skills. The Texas Learning Index (TLI) is an outcome metric, based on student performance on the TAAS, which allows for comparisons between administrations and between grades. The TLI has a common interpretation across grades: a score of 70 or above indicates the student performed at or above grade-level expectations. A student receiving the same score at consecutive grade levels made one year of academic progress. Analyses in the study were based on a transformation of the TLI that was conducted to induce normality.
STAR Math assessment	STAR Math is a computer-adaptive math test that assesses math skills. It combines computation and numeration items with word problems, estimation, statistics, charts and graphs, geometry, measurement, and algebra. STAR scores can appear as scaled scores or normal curve equivalent values.
Terra Nova mathematics subtest	The Terra Nova subtest is a national norm-referenced test that assesses academic performance in math.

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

Outcome measure	Study sample	Sample size (clusters/students)	Authors' findings from the study		WWC calculations			
			Mean outcome (standard deviation) ²		Mean difference ³ (Accelerated Math-comparison)	Effect size ⁴	Statistical significance ⁵ (at $\alpha = 0.05$)	Improvement index ⁶
			Accelerated Math group	Comparison group				
Nunnery & Ross, 2007 (quasi-experimental design)⁷								
2001 and 2002 transformed Texas Learning Index scores	Grade 8 cohort	4/992	1.21 ⁸ (0.47)	1.16 ⁸ (0.44)	0.05	0.11	ns	+4
Average for math achievement (Nunnery & Ross, 2007)⁹						0.11	ns	+4
Ysseldyke & Bolt, 2007 (randomized controlled trial with attrition)⁷								
2004 STAR Math normal curve equivalent scores	Grades 6–8	40/792	48.11 ¹⁰ (18.90)	44.45 ¹¹ (20.06)	3.66	0.19	ns	+7
2004 Terra Nova normal curve equivalent scores	Grades 6–8	40/851	46.89 ¹⁰ (18.67)	48.17 ¹¹ (18.69)	–1.28	–0.07	ns	–3
Average for math achievement (Ysseldyke & Bolt, 2007)⁹						0.06	ns	+2
Ysseldyke & Tardrew, 2007, grade 6 cohort (quasi-experimental design)⁷								
2002 STAR Math scale scores	Grade 6	17/326	773.43 ¹² (114.49)	762.80 ¹³ (93.82)	10.63	0.10	ns	+4
Average for math achievement (Ysseldyke & Tardrew, 2007 grade 6)⁹						0.10	ns	+4
Ysseldyke & Tardrew, 2007, grades 7 and 8 cohort (quasi-experimental design)⁷								
2002 STAR Math scale scores	Grades 7 and 8	8/149	801.14 ¹² (87.53)	786.47 ¹³ (83.33)	14.67	0.17	ns	+7
Average for math achievement (Ysseldyke & Tardrew, 2007, grades 7 and 8)⁹						0.17	ns	+7
Domain average for math achievement across all studies⁹						0.11	na	+4

ns = not statistically significant

na = not applicable

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. 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 [Technical Details of WWC-Conducted Computations](#).
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.

(continued)

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

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 level of statistical significance was reported by the study authors or, where 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 [Technical Details of WWC-Conducted Computations](#). In the case of Nunnery and Ross (2007), a correction for clustering was needed, so the significance levels may differ from those reported in the original study; no corrections for multiple comparisons were needed because there is only one outcome in this domain. In the case of Ysseldyke and Bolt (2007), a correction for clustering was needed, so the significance levels may differ from those reported in the original study; no corrections for multiple comparisons were needed because the WWC-computed effect sizes were not statistically significant. In the case of Ysseldyke and Tardrew (2007), a correction for clustering was needed, so the significance levels may differ from those reported in the original study; no corrections for multiple comparisons were needed because there is only one outcome in this domain.
8. Nunnery and Ross (2007, pp. 45–46) computed a transformation of the Texas Learning Index score to induce the distribution into normality to allow for an analysis of covariance.
9. 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.
10. The intervention group values from Ysseldyke and Bolt (2007) are the control group average plus the program coefficient from a regression analysis that controls for baseline pretest scores. The study authors provided the WWC with the program coefficient, unadjusted average, and standard deviations for both groups.
11. The control group average from Ysseldyke and Bolt (2007) are unadjusted.
12. The intervention group values from Ysseldyke and Tardrew (2007) are the comparison group means plus the difference in mean gains between the intervention (*Accelerated Math*) and comparison groups.
13. The comparison group means from Ysseldyke and Tardrew (2007) are unadjusted.

Appendix A4 Accelerated Math 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.¹

For the outcome domain of math achievement, the WWC rated *Accelerated Math* as no discernible effects. The remaining ratings (potentially negative effects and negative effects) were not considered, as *Accelerated Math* was assigned the highest applicable rating.

Rating received

No discernible effects: No affirmative evidence of effects.

- Criterion 1: No studies showing a statistically significant or substantively important effect, either *positive* or *negative*.

Met. No 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 three studies that evaluated math achievement and met WWC standards showed indeterminate effects.

Mixed effects: Evidence of inconsistent effects as demonstrated through EITHER of the following.

- 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 Intervention Rating Scheme](#).

Appendix A5 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence ¹
		Schools	Students	
Math achievement	3	>7 ²	≥2,259 ³	medium to large

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.”
2. Nunnery and Ross (2007) include four middle schools. Ysseldyke and Bolt (2007) include three middle schools. Ysseldyke and Tardrew (2007) do not report the number of middle schools.
3. Nunnery and Ross (2007) include 992 middle school students in the analysis sample. Ysseldyke and Tardrew (2007) include 326 grade 6 students and 149 grade 7 and 8 students. Ysseldyke and Bolt (2007) include 792 students in the analysis of the STAR Math outcome and 851 students in the analysis of the Terra Nova outcome, but the authors do not report the extent of overlap between the two samples.