What Works Clearinghouse

INSTITUTE OF EDUCATION SCIENCES

Elementary School Math September 28, 2006

Scott Foresman–Addison Wesley Elementary Mathematics

Program description

Scott Foresman–Addison Wesley Elementary Mathematics is a core curriculum for students at all ability levels in kindergarten through grade 6. The program supports students' understanding of key math concepts and skills and covers a range of mathematical content across grades. The curriculum focuses on questioning strategies, problem-solving skills, embedded assessment, and exercises tailored to students of different ability levels. It provides explicit problem-solving instruction, hands-on activities, and opportunities to extend students' mathematical understanding through reading and writing connections. According to its developer, *Scott Foresman–Addison Wesley Elementary Mathematics* is aligned to the National Council of Teachers of Mathematics Standards for the elementary grades.

Research

One study of *Scott Foresman–Addison Wesley Elementary Mathematics* met the What Works Clearinghouse (WWC) evidence standards. This study included more than 700 students in second and fourth grades, in a mix of urban, suburban, and rural schools in Washington, Wyoming, Virginia, and Kentucky.[1]

Effectiveness

Scott Foresman–Addison Wesley Elementary Mathematics was found to have no discernible effects on students' math achievement.

	Math achievement
Rating of effectiveness	No discernible effects
Improvement index[2]	Average: –2 percentile points Range: –7 to +3 percentile points

1. The evidence presented in this report is based on the available research. Findings and conclusions may change as new research becomes available.
2. These numbers show the average and range of improvement indices for all findings across the study.

Additional program information

Developer and contact

Developed by Pearson Scott Foresman, a division of Pearson Education, Inc., One Lake Street, Upper Saddle River, NJ 07458. Web: www.scottforesman.com. Telephone: (201) 236-7000.

Scope of use

The edition of *Scott Foresman–Addison Wesley Elementary Mathematics* reviewed in this report was published in 2004. Information is not available on the number or demographics of students, schools, or districts using this curriculum.

Teaching

Scott Foresman–Addison Wesley Elementary Mathematics consists of teacher-led lessons that follow a check-learn-check-practice sequence, emphasizing key math concepts and skills. Teachers check students' skills prior to each lesson, introduce the lesson, and then check students' understanding during the lesson. "Practice" sections in the text permit students to further demonstrate their understanding of concepts and apply this knowledge to solving real-life problems. Lessons (typically 45–60 minutes in length) are organized into chapters that extend over 2–8 weeks and use texts, workbooks, transparencies, manipulatives, and technology through group and individual activities.

Cost

The cost of *Scott Foresman–Addison Wesley Elementary Mathematics* varies based on the number of components included. For example, based on information from the developer, the costs in 2006 of some of the components used in the Resendez and Manley (2005) study for second grade are: student edition text ($28.60), homework workbook ($6.25), spiral review transparencies ($103), Problem of the Day transparencies/flipbook ($150), *Big Math Stories* (two volumes, $118), manipulative kits ($390), math game packs ($36), math vocabulary kit ($64.90), the Leveled Literature Library ($35), and test taking practice transparencies ($103).

Research

The WWC reviewed five studies of the *Scott Foresman–Addison Wesley Elementary Mathematics* program. One study (Resendez & Manley, 2005) was a randomized controlled trial that met WWC evidence standards. Thirty-five teachers in the second and fourth grades in six schools were randomly assigned to either the intervention condition (18 teachers) using *Scott Foresman–Addison Wesley Elementary Mathematics* or the comparison condition (17 teachers) using five distinct elementary math programs. The teachers in the intervention condition were in their first year of implementing the *Scott Foresman–Addison Wesley Elementary Mathematics* program. The comparison programs included curricula that took either a chapter-based, comprehensive basal approach or a strand/module-based investigative approach. The study compared math achievement outcomes of 375 students (200 second graders and 175 fourth graders) in the intervention condition with those of 344 students (188 second graders and 156 fourth graders) in the comparison condition. The four other studies did not meet WWC evidence screens.

Effectiveness

Findings

The WWC review of elementary school mathematics curriculum-based interventions addresses student outcomes in mathematics achievement.

Mathematics achievement. Resendez and Manley (2005) reported no statistically significant effects of the *Scott Foresman–Addison Wesley Elementary Mathematics* program on either the TerraNova Math Total Score or TerraNova Math Computation Score. The average effect across the two outcome measures was not large enough to be considered substantively important according to WWC criteria. Based on this study finding, the WWC rated the effect of *Scott Foresman–Addison Wesley Elementary Mathematics* on math achievement as indeterminate.

Effectiveness *(continued)*

Rating of effectiveness

The WWC rates interventions 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 (as calculated by the WWC[3]), the size of the difference between participants in the intervention condition and the comparison condition, and the consistency in findings across studies (see the WWC Intervention Rating Scheme). The WWC found Scott Foresman-Addison Wesley Elementary Mathematics to have no discernible effects on mathematics achievement.

The WWC found *Scott Foresman–Addison Wesley Elemenatry Mathematics* to have no discernible effects on mathematics achievement

Improvement index

For each outcome domain, the WWC computed an improvement index based on the effect size (see the Technical Details of WWC-Conducted Computations). The improvement index represents the difference between the percentile rank of the average student in the intervention condition versus 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 values denoting favorable results. The average improvement index for mathematics achievement is -2 percentile points, with a range of –7 to +3 percentile points across findings.

Summary

One study (Resendez & Manley, 2005) found indeterminate effects or, according to the WWC categorization scheme, "no discernible effects." The evidence presented in this report is limited and may change as new research emerges.

References

Met WWC evidence standards

Resendez, M., & Manley, M. A. (2005). *Final report: A study on the effectiveness of the 2004 Scott Foresman–Addison Wesley Elementary Math program.* Jackson, WY: PRES Associates, Inc.

Additional sources:

Resendez, M., & Sridharan, S. (2005). *Technical report: A study on the effectiveness of the 2004 Scott Foresman–Addison Wesley Elementary Math program.* Jackson, WY: PRES Associates, Inc.

Did not meet evidence screens

Gatti, G. G. (2004). Scott Foresman–Addison Wesley Math national effect size study. Available from Pearson Education, K–12 School Group, 1 Lake Street, Upper Saddle River, NJ 07458[4]

Klein, D. (2000). *High achievement in mathematics: Lessons from three Los Angeles elementary schools.* Washington, DC: Brookings Institution Press.[5]

Simpson, N. (2001). Scott Foresman California Mathematics validation study pretest-posttest results. (Report No. VM-17-3005-CA). Available from Pearson Scott Foresman, 1415 L Street, Suite 800, Sacramento, CA 95814[6]

WESTAT. (2003). *Analysis of field testing for Scott Foresman–Addison Wesley Mathematics 2004.* Rockville, MD: Author.[5]

For more information about specific studies and WWC calculations, please see the WWC *Scott Foresman–Addison Wesley Elementary Mathematics* Technical Appendices.

3. 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 see the WWC Tutorial on Mismatch. See the Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate the statistical significance. In the case of *Scott Foresman–Addison Wesley Elementary Mathematics* no corrections were needed.
4. Does not use a strong causal design: the study, which used a quasi-experimental design, did not establish that the comparison group was equivalent to the treatment group at the baseline in a pretest measure of math achievement.
5. Does not use a strong causal design: this is a qualitative study.
6. Does not use a strong causal design: the study did not use a comparison group.

Appendix

Appendix A1 Study characteristics: Resendez & Manley, 2005 (randomized controlled trial)

Characteristic	Description
Study citation	Resendez, M., and Manley, M. A. (2005). *Final report: A study on the effectiveness of the 2004 Scott Foresman-Addison Wesley Elementary Math program.* Jackson, WY: PRES Associates, Inc.
Participants	The participants in this study were second- and fourth-grade students. Ten classes of 200 second graders and eight classes of 175 fourth graders in six schools were randomly assigned to the intervention condition using *Scott Foresman–Addison Wesley Elementary Mathematics.* In the same six schools, nine classes of 188 second graders and eight classes of 156 fourth graders were assigned to the comparison condition using five distinct elementary math programs.
Setting	The six elementary schools were located in urban, suburban, and rural communities in Washington (one urban school), Wyoming (one rural and one suburban school), Virginia (one urban school), and Kentucky (two suburban schools).
Intervention	Students used the 2004 *Scott Foresman–Addison Wesley* (SFAW) text during the school year. This was the first year this program had been implemented in these schools. Teachers covered 70% (SD=15.3%) of the curriculum.
Comparison	Students used five distinct comprehensive math programs that took a basal or investigative approach and covered the same content as *Scott Foresman–Addison Wesley Elementary Mathematics.* Teachers covered 75% (SD=18.2%) of their math program.
Primary outcomes and measurement	The primary outcome measure was the TerraNova CTBS, Basic Multiple Assessment (Level 12, 2nd grade and level 14, 4th grade) with Plus Test. As cited in Resendez and Manley (2005), the TerraNova CTBS is a reliable and valid standardized test consisting of multiple choice, constructed response, and computational problems. According to the authors, it offers broad coverage of the mathematics content in most textbooks and reflects the National Council of Teachers of Mathematics (NCTM) standards. The TerraNova CTBS provides two overall scores from two separate tests: the TerraNova Math Total and TerraNova Math Computation Total. (See Appendix A2 for more detailed descriptions of outcome measures.)
Teacher training	Teachers in the intervention classrooms met with a SFAW professional trainer for a half-day session prior to implementing the curriculum in their classes. Two follow-up sessions of approximately two hours were conducted: one session occurred 4–8 weeks after the teachers began implementation, and a second session occurred 10–18 weeks after implementation. The training did not focus on professional development in the form of effective teaching strategies but instead focused on the vision of the program and how teachers could use the SFAW math program to help students make sense of mathematics.

Appendix A2 Outcome measures in the mathematics achievement domain

Outcome measure	Description
TerraNova CTBS Math Total score	The TerraNova CTBS Basic Multiple Assessment (Level 12, 2nd grade and Level 14, 4th grade) with Plus Test is a standardized test that, as described by Resendez and Manley (2005), was chosen for the study because of its validity, reliability, and sensitivity. The authors report that the test assesses content from the latest textbook series that are available from multiple publishers and reflects NCTM standards. Inter-rater reliability for the constructed response items was calculated by CTB McGraw Hill to range from 0.86 to 0.98 for all items on the second- and fourth-grade level tests (as cited in Resendez & Manley, 2005). The test provides two overall scores from two parts: The TerraNova Math Total and TerraNova Math Computation Total. The TerraNova Math Total (TNMT) score is based on multiple choice and constructed response for 43 Level 14 items and 34 Level 12 items. It is administered during two class sessions for a total of 75 minutes. The majority of items are word problems measuring basic, applied, and higher order thinking skills. TNMT Scores are reported as "scaled" and "normal curve equivalent" (as cited in Resendez & Manley, 2005).
TerraNova Math Computation Total score	The TerraNova Math Computation (TNMC) Total score is based on items in the CTBS-*Plus* test. It is a 20-item, multiple choice supplemental test that is administered for 20 minutes. The TNMC measures basic and advanced math computational skills (as cited in Resendez & Manley, 2005).

Appendix A3 Summary of study findings included in the rating for the mathematics achievement domain[1]

Outcome measure	Study sample	Sample size (students/ teachers)	Author's findings from the study: Mean outcome (standard deviation[2]) — *Scott Foresman–Addison Wesley* group[3]	Author's findings from the study: Mean outcome (standard deviation[2]) — Comparison group	WWC calculations: Mean difference[4] (*Scott Foresman–Addison Wesley* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Resendez & Manley, 2005 (randomized controlled trial)								
TerraNova Math Total score	Second- and fourth-grade students	645/35	55.59 (18.49)	54.14 (19.78)	1.45	0.08	ns	+3
TerraNova Math Computation score	Second- and fourth-grade students	533/35	53.89 (21.35)	57.49 (20.46)	–3.60	–0.17	ns	–7
Domain average[8] for mathematics achievement						–0.05	ns	–2

ns = not statistically significant

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices.
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 authors used multi-level modeling to explore the differences in outcomes. The intervention group means are the control means plus the program coefficient from the HLM analyses.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
5. For an explanation of the effect size calculation, please see the Technical Details of WWC-Conducted Computations.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The level of statistical significance was reported by the study author 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. See the Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of *Scott Foresman–Addison Wesley Elementary Mathematics,* no corrections were needed.
7. 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.
8. This row provides the study average, which in this instance is also the domain average. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The domain improvement index is calculated from the average effect size.

Appendix A4 *Scott Foresman–Addison Wesley Elementary Mathematics* rating for the math achievement domain

The WWC rates interventions as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.

For the outcome domain of math achievement, the WWC rated *Scott Foresman–Addison Wesley Elementary Mathematics* as having no discernible effects. It did not meet the criteria for positive effects because it had only one study, which had no statistically significant positive effects. Further, it did not meet the criteria for other ratings (potentially positive, mixed, potentially negative, or negative effects), because the study did not show statistically significant or substantively important effects, either positive or negative.

Rating received

No discernible effects: No affirmative evidence of effects.

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

  Met. The only study of *Scott Foresman–Addison Wesley Elementary Mathematics* showed an indeterminate effect.

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. Only one study of *Scott Foresman–Addison Wesley Elementary Mathematics* examined math achievement outcomes. Although it met WWC standards for a strong design, it did not show statistically significant effects.

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

  Met. The WWC analysis found no statistically significant or substantively important negative effects in this domain.

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. The WWC analysis found no statistically significant or substantively important positive effects in this domain.

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

  Not met. The only study of *Scott Foresman–Addison Wesley Elementary Mathematics* showed an indeterminate effect.

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. 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. The WWC analysis found no statistically significant or substantively important effects in this domain.

OR

(continued)

Appendix A4 *Scott Foresman–Addison Wesley Elementary Mathematics* rating for the math achievement domain *(continued)*

- 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. The WWC analysis found no statistically significant or substantively important effects in this domain.

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. No studies showing a statistically significant or substantively important *positive* effect. The number of studies showing statistically significant or substantively important *negative* effects is greater than the number showing statistically significant or substantively important *positive* effects.

  Not met. The WWC analysis found no statistically significant or substantively important negative effects in this domain.

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 is based on a strong design.

  Not met. The only study of *Scott Foresman–Addison Wesley Elementary Mathematics* showed an indeterminate effect.

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

  Met. The WWC analysis found no statistically significant or substantively important positive effects in this domain.