What Works Clearinghouse™

ies INSTITUTE OF EDUCATION SCIENCES

Elementary School Math January 2013

enVisionMATH

Program Description[1]

enVisionMATH, published by Pearson Education, Inc., is a core curriculum for students in kindergarten through grade 6. The program seeks to help students develop an understanding of math concepts through problem-based instruction, small-group interaction, and visual learning with a focus on reasoning and modeling. Differentiated instruction and ongoing assessment are used to meet the needs of students at all ability levels.

Research[2]

The What Works Clearinghouse (WWC) identified one study of *enVisionMATH* that both falls within the scope of the Elementary School Mathematics topic area and meets WWC evidence standards. The study meets WWC evidence standards without reservations, and included 1,156 elementary school students in the second and fourth grades in eight locations across the United States.[3]

The WWC considers the extent of evidence for *enVisionMath* on the math performance of elementary school students to be small for the mathematics achievement domain, the only outcome domain examined for studies reviewed under the Elementary School Mathematics topic area. (See the Effectiveness Summary on p. 4 for further description of this domain.)

Report Contents

Effectiveness

enVisionMath was found to have potentially positive effects on mathematics achievement for elementary school students.

Table 1. Summary of findings[4]

Outcome domain	Rating of effectiveness	Improvement index (percentile points) Average	Improvement index (percentile points) Range	Number of studies	Number of students[3]	Extent of evidence
Mathematics achievement	Potentially positive effects	+6	+1 to +9	1	1,156	Small

Program Information

Background

enVisionMATH was developed and is distributed by Pearson Education, Inc. Address: One Lake Street, Upper Saddle River, NJ 07458. Web: www.pearsonschool.com. Telephone: (201) 236-7000.

Program details

enVisionMATH consists of 120–130 teacher-led lessons for each grade. Lessons are designed to be completed at the pace of one per day. Each lesson includes daily review and a small-group, problem-based activity, followed by guided and independent practice activities. Instructors use daily assessments to track student progress. These assessments also allow for targeting of additional practice and homework activities for students needing more support. Lessons are organized into a customizable sequence of topics and use texts, workbooks, manipulatives, and technology, incorporating both group and individual activities.

Cost

The student editions of *enVisionMATH* for kindergarten through grade 2 cost $26.97 per student per year, with discounts available to districts through a subscription model. This cost includes access to interactive digital courseware for one year. The student editions for grades 3–6 cost $65.97 and include access to the interactive digital courseware for six years. The teacher's edition for each grade costs $525. Additional materials, including workbooks, manipulatives, digital resources, instructional materials, and teacher guides may be purchased separately, with prices varying by material and quantity purchased.

Research Summary

The WWC identified two studies that investigated the effects of *enVisionMATH* on the math performance of elementary school students.

The WWC reviewed the two studies against group design evidence standards. One of the studies (Resendez & Azin, 2008) is a randomized controlled trial that meets WWC evidence standards without reservations and is summarized in this report. The other study does not meet WWC evidence standards. Citations for both studies are in the References section, which begins on p. 5.

Table 2. Scope of reviewed research[5]

Grade	2, 4
Delivery method	Whole class
Program type	Curriculum
Studies reviewed	2 studies
Group design studies that meet WWC evidence standards	
• without reservations	1 studies
• with reservations	0 studies

Summary of study meeting WWC evidence standards without reservations

Resendez and Azin (2008) examined the effects of *enVisionMATH* on 1,156 second- and fourth-grade students in eight elementary schools located in eight states.[6] Within each school, teachers in grades 2–5 were randomly assigned to teach either *enVisionMATH* or the existing math curriculum in place at their school.[3] The eight study schools each used a different comparison curricula.

Summary of studies meeting WWC evidence standards with reservations

No studies of *enVisionMATH* meet WWC evidence standards with reservations.

Effectiveness Summary

The WWC review of *enVisionMATH* for the Elementary School Mathematics topic includes student outcomes in one domain: mathematics achievement. The findings below present the authors' estimates and WWC-calculated estimates of the size and statistical significance of the effects of *enVisionMATH* on the math performance of elementary school students. For a more detailed description of the rating of effectiveness and extent of evidence criteria, see the WWC Rating Criteria on p. 12.

Summary of effectiveness for the mathematics achievement domain

One study reported findings in the mathematics achievement domain.

Resendez & Azin (2008) found, and the WWC confirmed, three positive and statistically significant differences between students in the *enVisionMATH* and comparison groups on mathematics achievement outcomes, and one difference that was not statistically significant. Based on the statistically significant findings, the WWC characterizes this study as having statistically significant positive effects.

Thus, for the mathematics achievement domain, there was one study showing a statistically significant positive effect, with no studies showing a statistically significant or substantively important negative effect or an indeterminate effect. This results in a rating of potentially positive effects, with a small extent of evidence.

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

Rating of effectiveness	Criteria met
Potentially positive effects *Evidence of a positive effect with no overriding contrary evidence.*	In the study that reported findings, the estimated impact of the intervention on outcomes in the *mathematics achievement domain* was a statistically significant positive effect.
Extent of evidence	**Criteria met**
Small	One study that included 1,156 students in eight schools reported evidence of effectiveness in the *mathematics achievement domain*.[3]

References

Study that meets WWC evidence standards without reservations

Resendez, M., & Azin, M. (2008). *A study on the effects of Pearson's 2009 enVisionMATH program. 2007–2008: First year report.* Jackson, WY: PRES Associates, Inc.

Additional source:

Resendez, M., Azin, M., & Strobel, A. (2009). *A study on the effects of Pearson's 2009 enVisionMATH program: Final summative report.* Jackson, WY: PRES Associates, Inc.

Studies that meet WWC evidence standards with reservations

None

Study that does not meet WWC evidence standards

Resendez, M., & Azin, M. (2010). *A study on the relationship between Pearson's 2009 enVisionMATH program and student math performance among English language learners, minorities, and economically disadvantaged students: Special report.* Jackson, WY: PRES Associates, Inc. Retrieved from: http://www.pearsoned.com/wp-content/uploads/envisionmath-archival-study-on-minorities-and-ells-final.pdf. 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.

Studies that are ineligible for review using the Elementary School Mathematics Evidence Review Protocol

None

Appendix A: Research details for Resendez & Azin, 2008

Resendez, M., & Azin, M. (2008). *A study on the effects of Pearson's 2009 enVisionMATH program. 2007–2008: First year report.* Jackson, WY: PRES Associates, Inc.

Table A. Summary of findings

Meets WWC evidence standards without reservations

Outcome domain	Sample size	Study findings	
		Average improvement index *(percentile points)*	Statistically significant
Mathematics achievement	8 schools/1,156 students[3]	+6	Yes

Setting

The study was conducted in eight elementary schools in eight states across the United States: Colorado, Kentucky, Massachusetts, Montana, New Hampshire, North Carolina, Ohio, and Tennessee.

Study sample

Teachers of grades 2–5 were randomly assigned within participating schools either to an intervention condition in which teachers used the *enVisionMATH* curriculum or to a comparison condition in which teachers continued to use the elementary school mathematics curriculum already in place at their school. Schools in the study were chosen based on geographic variation, the presence of multiple teachers in each grade (grades 2–5), a historical pattern of low student mobility rates, expressed interest in participating in this study, and use of a curriculum that the researchers felt was distinct from *enVisionMATH*. The study's analysis of outcomes following one year of exposure to the intervention includes students in grades 2 and 4.[6] A total of 54 teachers and 1,156 students were included in the analysis. Based on a partial sample at baseline, which is larger than the analysis sample size, the intervention group was approximately 89% White and 54% male, and the comparison group was approximately 95% White and 52% male.

Intervention group

enVisionMATH is a core elementary mathematics curriculum that emphasizes interactive learning with frequent assessments to support individualized instruction. Intervention teachers were provided with performance guidelines, workbooks, and assessments to implement the *enVisionMATH* curriculum.

Comparison group

Teachers in the comparison condition used the elementary school mathematics curriculum already in place at their school. Study authors selected the sample of participating schools based on the schools' existing curricula being distinct, in their estimation, from *enVisionMATH*. The eight study schools each used a different comparison curriculum.

Outcomes and measurement

Eligible outcomes included two subtests of the Metropolitan Achievement Test (MAT 8), Form V (Math Computation; Math Concepts and Problem-solving), the Concepts and Communication subtest of the Group Mathematics Assessment and Diagnostic Evaluation (GMADE), and an open-ended assessment of problem solving and reasoning skills. This latter assessment is based on the standard version of the Balanced Assessment in Mathematics for fourth-grade students and a nonstandard version developed by the study authors for second-grade students. Outcome data were collected at the beginning and end of the school year. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

Teachers were trained by a Pearson staff member for about 5–6 hours before the *enVision-MATH* curriculum was used in class. The trainers were given guidelines on teacher trainings and instructed to provide an overview of the components of *enVisionMATH*, then model examples for teachers. A three-hour follow-up training was given within 2–3 weeks of the start of the school year, and several sites received one or two additional 3–4 hour follow-up training sessions.

Appendix B: Outcome measures for each domain

Mathematics achievement	
Group Mathematics Assessment and Diagnostic Evaluation (GMADE): Concepts and Communication subtest	The GMADE, published by Pearson Assessment, is a norm-referenced standards-based assessment of mathematical skills. The Concepts and Communication subtest uses pictures, symbols, words, and numbers to address language, vocabulary, and representational aspects of mathematics (as cited in Resendez & Azin, 2008).
Metropolitan Achievement Test (MAT 8): Math Computation subtest	The MAT 8, Form V, published by Harcourt Assessment, is a norm-referenced standardized mathematics assessment. Using applied problems, it assesses mathematical knowledge, problem solving, communication, and mathematical reasoning. The Math Computation subtest examines students' ability to complete arithmetic operations, which depending on the level may include addition, subtraction, multiplication, and division of different types of numbers (whole numbers, decimals, fractions, and percents) (as cited in Resendez & Azin, 2008).
MAT 8: Math Concepts and Problem-solving subtest	The MAT 8, Form V, published by Harcourt Assessment, is a norm-referenced standardized mathematics assessment. Using applied problems, it assesses mathematical knowledge, problem solving, communication, and mathematical reasoning. The Math Concepts and Problem-solving subtest assesses students' ability to apply mathematical reasoning to different types of problems. It includes the concepts of numbers and operation, patterns and relationships, geometry, measurement, data and probability, and estimation (as cited in Resendez & Azin, 2008).
Open-Ended Assessment of Problem Solving and Reasoning Skills	The fourth-grade version of this measure is the Balanced Assessment in Mathematics (BAM), published by CTB. BAM uses multi-step tasks to demonstrate mathematical performance. No second-grade version of BAM exists, so the authors created their own version, selecting or adapting items from two major sources: the Elementary Grades Assessment and the National Council of Teachers of Mathematics' (NCTM) Mathematics Assessment Sampler (as cited in Resendez & Azin, 2008). The psychometric properties of the nonstandard second-grade version of the assessment, which the study authors provided to the WWC, meet the standards required by the Elementary School Mathematics review protocol. This report refers to the combination of standard BAM and the author-developed measure as an open-ended assessment of problem solving and reasoning skills.

Appendix C: Findings included in the rating for the mathematics achievement domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Resendez & Azin, 2008[a]								
GMADE: Concepts and Communication subtest	Grades 2 and 4	8 schools/ 1,118 students	99.55 (10.25)	97.21 (10.59)	2.34	0.22	+9	0.01
MAT 8: Math Computation subtest	Grades 2 and 4	8 schools/ 1,156 students	645.83 (60.54)	636.04 (60.65)	9.79	0.16	+6	0.02
MAT 8: Math Concepts and Problem-solving subtest	Grades 2 and 4	8 schools/ 1,154 students	636.84 (51.06)	634.98 (52.24)	1.86	0.04	+1	0.52
Open-Ended Assessment of Problem Solving and Reasoning Skills	Grades 2 and 4	8 schools/ 1,151 students	75.20 (17.57)	72.05 (17.65)	3.15	0.18	+7	0.01
Domain average for mathematics achievement (Resendez & Azin, 2008)						**0.15**	**+6**	**Statistically significant**
Domain average for mathematics achievement across all studies						**0.15**	**+6**	**na**

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the average change expected for all students who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average student's percentile rank that can be expected if the student is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of the study's domain average was determined by the WWC. MAT 8 = Metropolitan Achievement Test. GMADE = Group Mathematics Assessment and Diagnostic Evaluation. na = not applicable.

[a] For Resendez & Azin (2008), a correction for multiple comparisons was needed but did not affect significance levels. The *p*-values presented here were reported in the original study. The authors' analysis utilized hierarchical linear modeling (HLM), which accounts for the nesting of cases within clusters (in this case, students within classrooms). This approach obviates the need for a clustering correction, which might otherwise be needed given the teacher-level random assignment. The intervention group value is the comparison group mean plus the program coefficient from the HLM analysis. Sample sizes in the table were obtained from the Technical Report (provided by the authors) for Resendez & Azin (2008).

Appendix D: Description of supplemental findings for Year 2 for the mathematics achievement domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Resendez & Azin, 2009[a]								
MAT 8: Math Concepts and Problem-Solving subtest	Grades 3 and 5	6 schools/ 681 students	684.57 (60.34)	673.14 (50.51)	11.43	0.21	+8	0.00

Table Notes: The supplemental findings presented in this table are additional findings from the studies in this report that do not factor into the determination of the intervention rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the average change expected for all students who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average student's percentile rank that can be expected if the student is given the intervention. MAT 8 = Metropolitan Achievement Test.

[a] For Resendez & Azin (2009), no corrections for clustering or multiple comparisons were needed. The *p*-values presented here were reported in the original study. The study's Year 2 analysis included four measures of mathematics achievement; however, only the MAT 8: Math Concepts and Problem-Solving subtest meets WWC standards without reservations. The authors' analysis utilized HLM, which accounts for the nesting of cases within clusters (in this case, students within classrooms). This approach obviates the need for a clustering correction, which might otherwise be needed given the teacher-level random assignment. The intervention group value is the comparison group mean plus the program coefficient from the HLM analysis. This study also reported findings for additional measures of mathematics achievement; however, these findings were based on analysis that exhibited high attrition and failed to demonstrate baseline equivalence of the research groups. Therefore, findings related to these additional measures do not meet WWC standards. Sample sizes in the table were obtained from the Technical Report (provided by the authors) for Resendez & Azin (2009).

Endnotes

[1] The descriptive information for this program was obtained from a publicly available source: the developer's website (http://www.pearsonschool.com, downloaded January 2012). The WWC requests developers review the program description sections for accuracy from their perspective. The program description was provided to the developer in January 2012, and the WWC incorporated feedback from the developer. 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 December 2011.

[2] The studies in this report were reviewed using the Evidence Standards from the WWC Procedures and Standards Handbook (version 2.1), along with those described in the Elementary School Mathematics review protocol (version 2.0). The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

[3] The number of students included in the analysis in Resendez & Azin (2008) varied by the outcome measure examined. When referencing the number of students included in the study, we use the maximum number of students included across all the outcome measures. For sample sizes for each outcome measure, see Appendix C.

[4] For criteria used in the determination of the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on page 12. These improvement index numbers show the average and range of student-level improvement indices for all findings across the studies.

[5] Grade, delivery method, and program type refer to the studies that meet WWC evidence standards without or with reservations.

[6] The intervention rating in this report focuses on findings from the stronger research design, which was from the study authors' analysis of the first year of exposure to the intervention for students in grades 2 and 4. The study authors also examined the same students in six of the eight sample schools following a second year of exposure to the intervention (two schools adopted new math curricula school-wide for 2008–09 and did not participate in the second year of the study). The intervention rating focuses on the first year because findings for all four measures of mathematics achievement from the analysis of the first year meet WWC standards without reservations, whereas only one of four measures of mathematics achievement from the analysis of the second year meets WWC standards. Findings for three of the four measures of mathematics achievement in the study's second year do not meet WWC standards because they have high attrition and there are large differences in the baseline characteristics of the intervention and comparison groups. Findings from the one measure of mathematics achievement in the study's second year that does meet WWC standards are presented in Appendix D.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2013, January). *Elementary School Mathematics intervention report: enVisionMATH*. Retrieved from http://whatworks.ed.gov.

WWC Rating Criteria

Criteria used to determine the rating of a study

Study rating	Criteria
Meets WWC evidence standards without reservations	A study that provides strong evidence for an intervention's effectiveness, such as a well-implemented RCT.
Meets WWC evidence standards with reservations	A study that provides weaker evidence for an intervention's effectiveness, such as a QED or an RCT with high attrition that has established equivalence of the analytic samples.

Criteria used to determine the rating of effectiveness for an intervention

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

Criteria used to determine the extent of evidence for an intervention

Extent of evidence	Criteria
Medium to large	The domain includes more than one study, AND The domain includes more than one school, AND The domain findings are based on a total sample size of at least 350 students, OR, assuming 25 students in a class, a total of at least 14 classrooms across studies.
Small	The domain includes only one study, OR The domain includes only one school, OR The domain findings are based on a total sample size of fewer than 350 students, AND, assuming 25 students in a class, a total of fewer than 14 classrooms across studies.

Glossary of Terms

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

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

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

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

Domain A domain is a group of closely related outcomes.

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

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

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

Extent of evidence An indication of how much evidence supports the findings. The criteria for the extent of evidence levels are given in the WWC Rating Criteria on p. 12.

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

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

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

Randomized controlled trial (RCT) A randomized controlled trial (RCT) is an experiment in which investigators randomly assign eligible participants into intervention and comparison groups.

Rating of effectiveness The WWC rates the effects of an intervention in each domain based on the quality of the research design and the magnitude, statistical significance, and consistency in findings. The criteria for the ratings of effectiveness are given in the WWC Rating Criteria on p. 12.

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

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

Statistical significance Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < 0.05$).

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

Please see the WWC Procedures and Standards Handbook (version 2.1) for additional details.