2007-08: First Year Report

A Study on the Effects of Pearson's 2009 enVisionMATH Program

PREPARED BY:

MIRIAM RESENDEZ, M.A., SENIOR RESEARCHER
MARIAM AZIN, PH.D., PRESIDENT

SEPTEMBER, 2008

For inquiries, please contact PRES Associates at:
info@presassociates.com
(307) 733-3255

Executive Summary

Research suggests that children's performance at the end of elementary school is an important predictor of their future educational success (National Research Council, 2001). However, research also continues to show that U.S. students are not being adequately prepared to meet the demands of careers in a global economy, including advanced skills in critical thinking and mathematics. In order to more fully prepare students' with the skills they need to become successful in higher level math courses, as well as their futures, Pearson released enVisionMATH (2009), a comprehensive elementary math program. This research-based program aims to provide problem-based interactive learning opportunities that are enhanced by visual learning strategies to create more in-depth connections for students.

In order to determine the effectiveness of enVisionMATH in helping students attain critical math skills and understanding, Planning, Research, and Evaluation Services (PRES) Associates, Inc. is conducting a two year study at the 2nd through 5th grades. This randomized control trial (RCT), which commenced in the Fall of 2007, was conducted in the 2nd and 4th grades during the 2007-2008 school year and will continue during the 2008-2009 school year in the 3rd and 5th grades. This report presents findings from the first year (2007-2008) of the RCT.

For the first year, the final sample consisted of 1197 students spread across eight geographically-dispersed schools. Teachers were randomly assigned to treatment (n=29) and control conditions (n=30).

Major findings, organized by the key evaluation questions, include:

Does math ability improve over the course of participating in enVisionMATH? Does this vary by different types of students and levels of implementation?

Students using enVisionMATH significantly improved over the course of the school year in the areas of math concepts and problem-solving, math computation, math vocabulary, and communication in math. Specifically, results showed that enVisionMATH students demonstrated significant percentile gains of:

- 19 points on concepts and problem-solving,
- 33 points on computation,
- 15 points on math vocabulary, and
- 36 points on math communication.

Another way to look at norm-referenced assessment results is to look at the percentile rankings of students relative to a national sample. It is a general rule of thumb that if a student makes a year's growth for a year of instruction, then the percentile rank will remain the same. enVisionMATH students had higher percentile rankings at post-testing than pre-testing on the MAT8 Computational subtest (54th to 58th percentile) and the GMADE (36th to 45th percentile). This means that enVisionMATH students grew more than would be expected over the course of a typical academic year as measured by the math computation and vocabulary tests.

Furthermore, the enVisionMATH program worked just as well with 2nd and 4th graders, females and males, White and non-White students, special education and non-special education students, students receiving free/reduced lunch and those not receiving this aid, and students at various math levels. Although a greater rate of improvement was demonstrated for certain subgroups of students, enVisionMATH

students in all subpopulations showed significant learning gains on all assessment measures.

Since there was some variation observed in overall implementation of the enVisionMATH program among treatment teachers, analyses were performed to examine if this affected student performance. Results showed that there was a significant relationship between overall enVisionMATH implementation levels and improvement in math performance. Specifically, preliminary analyses showed that students whose teachers implemented the major components of enVisionMATH with high fidelity showed greater improvement as compared to students of teachers who implemented enVisionMATH with low fidelity and did not use the major program components on a regular basis.

How does mathematics performance differ between students who use enVisionMATH as compared to students who do not use this program? Do effects on math performance differ across types of students or settings?

Results showed positive effects of the enVisionMATH program. Elementary students who used enVisionMATH in 2007-08 showed greater gains in math computation, math vocabulary, and math problem-solving and communication as compared to students who used other math programs. In addition, while effects were small (ranged from .20 to .24), they are both significant and meaningful – especially given the short duration of the study and the applied settings in which the study was undertaken. Given that this study will continue into a second year (2008-2009 school year), larger effects could be expected if such trends continue and students are exposed to enVisionMATH for even longer periods of time.

Analysis of subgroup differences also showed significant effects. Notably, the following types of enVisionMATH students showed significantly greater gains in math performance as compared to control students in these subgroups:

- 4th graders
- Minorities
- Females
- High math ability students

Moreover, of the 16 significant subgroup differences observed, 14 were in favor of the enVisionMATH program (i.e., enVisionMATH students outperformed control students).

In addition, the positive effects obtained on the enVisionMATH program were observed across a number of different schools who used a variety of types of control programs. Specifically, enVisionMATH students performed significantly better than control students who used programs that were purely investigative and inquiry-based as well as students who used more traditional basal math programs. This consistency in findings across different curricula, schools, and measures lends credence to the conclusion that enVisionMATH positively impacts student math knowledge and skills.

Does participation in enVisionMATH result in other positive student outcomes (e.g., positive attitudes towards math, etc.)?

While the main focus of the enVisionMATH program is to improve upon important math skills and understanding, other measures were included to explore if enVisionMATH was associated with positive impacts on student and teacher attitudes, and classroom practices. Results showed that enVisionMATH students enjoyed math more and perceived greater teacher support as compared to control students.

enVisionMATH and control students had similar positive attitudes in various other areas (e.g., importance of math, attitudes about math success, and so forth).

Results also showed that enVisionMATH had positive effects on teacher attitudes and instructional practices. In particular, enVisionMATH teachers noted that they were more prepared to carry out various mathematics activities, and in fact, tended to engage in a greater variety of mathematics activities and strategies as compared to control teachers. enVisionMATH teachers also reported greater confidence in their ability to teach math as compared to control teachers and showed a significant increase in their perceived knowledge of NCTM standards and focal points. In addition, following one year of enVisionMATH use, teachers tended to be more likely to incorporate inquiry-based strategies into their math instruction. That said, the overall pedagogical approach employed by enVisionMATH teachers tended to blend inquiry-based and traditional approaches. This is consistent with the overall enVisionMATH philosophy in that it attempts to blend both types of pedagogies.

What did users of enVisionMATH think about the program?

The enVisionMATH program was highly regarded by the vast majority of teachers. A full 92% of enVisionMATH teachers surveyed indicated that the program was an effective tool for mathematics instruction. In addition, 89% of enVisionMATH teachers indicated that they would recommend their math program in comparison to only 39% of control teachers. Comparisons also revealed that enVisionMATH teachers rated their math programs' resources (e.g., professional development embedded within program, reading/writing in math activities, technology resources, review materials, etc) as more useful compared to control teachers.

Teachers noted a variety of specific program components when asked to identify the three things they liked best about the enVisionMATH program. However, a few items emerged as favorites, including:

- The amount of story problems and the "thinking about math" involved in the program.
- The Daily Spiral Review
- The visual representation of concepts throughout the program
- The manipulatives provided and hands-on aspects of the program
- The amount of writing provided in the program
- The design of the pouches and Teacher's Edition
- The technology and online resources provided

Students who used enVisionMATH also enjoyed the program. Significant differences were observed in that enVisionMATH students rated their math program higher than control students; for example, 73% of enVisionMATH students noted that they liked the program used in math class as compared to 62.8% of control students.

In summary, results found from the first year of the RCT show that the enVisionMATH program does have a significant effect on student math performance as compared to students who use other math programs, including purely inquiry-based and traditional basal math programs. enVisionMATH students showed significantly better performance than control students across portions of all three national math assessments used as outcome measures in this study. Specifically, students who used enVisionMATH outperformed students who used other types of math programs in the areas of math computation, math vocabulary, and math problem-solving and communication. While these main findings can be classified as small effects, it should

be noted that such small effects are typical of applied research, especially curricular research involving comparisons across core curricula. In addition, the significant and consistent positive effects of enVisionMATH are noteworthy given that this was the first year of a two year evaluation. Indeed, if the current pattern of results hold, larger effects can be expected following two years of enVisionMATH as teachers become more familiar with the program and its resources, and treatment students have prolonged exposure to the program.

Table of Contents

Table of Figures and Tables

Figures

Tables

Project Background

> ***"During most of the 20th century, the United States possessed peerless mathematical prowess—not just as measured by the depth and number of the mathematical specialists who practiced here but also by the scale and quality of its engineering, science, and financial leadership, and even by the extent of mathematical education in its broad population. But without substantial and sustained changes to its educational system, the United States will relinquish its leadership in the 21st century."***
>
> ***--Report of the National Mathematics Advisory Panel***

Focus on the education of U.S. students has recently shifted from reading/language arts to mathematics. This renewed focus on improving upon the mathematics skills of students is due, in part, to research that continues to show that U.S. students are not being adequately prepared to meet the demands of future careers in a global economy, including advanced skills in critical thinking and mathematics. While the latest results from the National Assessment of Education Progress (2007) points to improvements in the math performance of 4th and 8th graders, international comparisons have shown that U.S. students are falling behind in math as compared to students of other countries (Mullis, Martin & Foy, 2005).

Evidence of the urgency needed to adequately prepare students to succeed in higher level math courses, and ultimately their careers, can be seen in the presidents' call for the creation of a National Math Panel in April 2006. This panel was given the charge to use the best scientific research available and "…to foster greater knowledge of and improved performance in mathematics among American students" (Report of the National Mathematics Advisory Panel, 2008). While a major focus of the National Math Panel was centered on defining and improving important Algebra skills, there is evidence that children's performance at the *end of elementary school* is an important predictor of their future educational success (National Research Council, 2001). As so aptly stated by the National Association for the Education of Young Children: "If progress in improving the mathematics proficiency of Americans is to continue, much greater attention must be given to *early* mathematics experiences" (NAEYC, 2008).

> ***"We must teach number and math concepts early, we must help students believe they can improve their math skills and we must ensure they fully comprehend algebra concepts by the time they graduate from high school."***
>
> ***—U.S. Secretary of Education Margaret Spellings***

The 2009 Pearson enVisionMATH program, aligned to curriculum focal points suggested by the National Council of Teachers of Mathematics (NCTM), is a new elementary math program that promises to be an effective instructional program for elementary school students. This research-based program aims to provide problem-based interactive learning opportunities that are enhanced by visual learning strategies to create more in-depth connections for students. The program incorporates a blended approach of traditional and investigative learning that emphasizes embedded assessment and data driven remediation. Given how important math skills are to the future success of children throughout their lifetime, programs that can help in the development of such skills need

to be looked at carefully to determine the extent to which they help students attain such skills. Indeed, the No Child Left Behind Act of 2001(NCLB) mandates that educational materials purchased with public funds be proven by scientific research to improve student achievement in the classroom.

Planning, Research, and Evaluation Services (PRES Associates), an external, independent, educational research firm with nearly 20 years of experience in applied educational research and evaluation is conducting a two-year study to examine the effectiveness of the 2009 Pearson enVisionMATH program in helping elementary students improve their mathematics skills and understanding. This randomized control trial (RCT), which commenced in the Fall of 2007, was conducted in the 2nd and 4th grades during the 2007-2008 school year and will continue during the 2008-2009 school year in the 3rd and 5th grade. This report presents findings from the first year (2007-2008) of the RCT.

Project Overview

The overarching purpose of this study is to rigorously evaluate the effectiveness of the 2009 enVisionMATH program in helping elementary students attain critical math skills. Specifically, this study is designed to address the following research questions:

1. Does math ability improve over the course of participating in enVisionMATH? Does this vary by different types of students and levels of implementation?

2. How does mathematics performance differ between students who use enVisionMATH as compared to students who do not use this program?

3. Do effects of enVisionMATH on student achievement vary as a function of different student or school level characteristics? That is, do study findings vary across different types of students, and settings?

4. Does participation in enVisionMATH result in other positive student outcomes (e.g., positive attitudes towards math, etc.)?

5. What do users of enVisionMATH think about the program? What aspects of the program do they find most useful? Least useful? What, if any, suggestions for program improvement do they have?

This report presents descriptive information and results of the first year of the RCT. Specifically, the remainder of this report includes: 1) a description of the design and methodology; 2) sample and site information, including descriptions of enVisionMATH implementation; 3) results of the first year of the evaluation; and 4) conclusions and a summary of activities that will occur during the second year of the evaluation. In addition, an accompanying Technical Report presents detailed statistical results of all baseline, attrition and assessment analyses conducted on the first year data, including the analytical goals and framework employed.

Design and Methodology

Research Design

The present study was designed to address all standards and criteria described in the What Works Clearinghouse (WWC) Study Review Standards (2008) and the Joint Committee on Standards for Educational Evaluation's Program

Evaluation Standards (1994). Appendix A outlines how this study addresses each of the WWC Study Review standards that help ensure the quality of scientifically-based research.

The research design consists of a two-year randomized control trial, with random assignment of teachers to a treatment (i.e., use of enVisionMATH) or control group (i.e., use of other elementary math program) *within* schools[1]. Other important design and methodological features include:

- The study was conducted in the 2nd and 4th grades during the 2007-2008 school year and will extend into the 3rd and 5th grades during the 2008-2009 school year.
- Random assignment occurred at the teacher level. Teachers at all grade levels (2nd to 5th) were assigned to treatment or control conditions at the beginning of the study.
- Clear site selection criteria were established along with accompanying rationale.
- To the extent possible, the control programs to which enVisionMATH was compared were selected to be as distinct as possible given the common content taught.
- Extensive background data[2] was collected on instructional activities and materials employed in both treatment and control conditions so that distinctive pedagogical elements could be described given the common content taught.
- The threat of differential attrition was addressed via: 1) the initial site selection process[3]; 2) random assignment within schools, at the teacher level, to help ensure that attrition is relatively constant across both treatment and control groups; and 3) the characteristics of students who dropped out were statistically compared between treatment and control groups.
- Extensive implementation guidelines and monitoring procedures[4] were embedded to ensure the fidelity of treatment implementation.
- A battery of assessments aligned to national math standards and offering a broad-range of content matter and item types was used in order to enhance the sensitivity of the study to picking up treatment effects.
- The study employed pre/post measures of, among other things, (1) student performance; (2) school, teacher and math-related attitudes; (3) teacher practices; and (4) teacher knowledge and characteristics.
- Student assessments, surveys, and classroom observation forms are valid and reliable as shown by technical documentation and statistical analyses performed.
- The study employed the use of statistical controls as well as random assignment to establish initial group equivalence[5].
- Analyses of assessment data were primarily conducted via multilevel modeling (MLM) with student and teacher level data to take into account dependency issues. In addition, the teacher level of analysis used in MLM matches the unit of random assignment.

[1] There are a number of reasons why random assignment to treatment conditions was done at the teacher level. The most important reason for selecting this level of assignment is that such a design provides an opportunity to help establish *causality* by eliminating the threat that school level factors could have potentially contributed to differences between treatment and control groups. An important issue to be considered with this design option, however, is that procedures must be put into place to ensure that the treatment and control classes are not contaminated through teachers sharing of enVisionMATH materials. Indeed, this was accomplished through stringent guidelines provided to the teachers and close monitoring of their instruction and use of resources by researchers.

[2] Descriptive information was obtained so that, even if not all extraneous variables related to the outcome measures can be controlled, they can at least be measured and used as covariates in subsequent analyses.

[3] Sites that historically had more than 20% student attrition were not used in the study.

[4] Training provided and implementation guidelines reflect how the enVisionMATH program should typically be used in schools.

[5] Random assignment helps to create group equivalence. However, it must be noted that with small sample sizes random assignment in and of itself does not assure initial group equivalence (Lipsey, 1990).

Table 1 displays the timeline for the important study activities during the first year of the RCT. More detailed information on these activities, as well as measures being used are discussed in the following section.

Measures

This section reviews the outcome and assessment measures that were administered, including descriptions of the items, and available reliability and validity information.

ASSESSMENT

In order to enhance the sensitivity of the RCT to detect any effects associated with enVisionMATH, a battery of outcome measures were selected. Assessment selection was based on a thorough literature review of existing assessments to identify tests that were valid, reliable, measured various mathematics skills (e.g., computational skills, math vocabulary, problem-solving, etc.), and that included content that reflected important concepts and skills in major math textbook series, mathematics literature, and the NCTM Curriculum and Evaluation Standards and subsequent NCTM publications. Student assessments consist of the following:

- *Metropolitan Achievement Test (MAT8)*: Norm-referenced standardized assessment that allows for an in-depth examination of the following two math areas: (1) Math Concepts and Problem-Solving and (2) Math Computation. All items are multiple-choice.
- *GMADE – Concepts & Communication* subtest: This subtest uniquely addresses the language, vocabulary, and representations of mathematics. All items are multiple-choice.
- *Balanced Assessment of Mathematics (BAM)*: This test adds an open-ended, performance-based component so that students can communicate what they know and allows for the assessment of those students who are better at communicating knowledge. It specifically targets problem-solving and reasoning skills.

Metropolitan Achievement Test. The MAT8 Form V, published by Harcourt Assessment, is a group-administered, norm-referenced test that measures student achievement for students in grades K-12. According to the publisher, the mathematics tests assess appropriate mathematics content and process skills at every level, grounding questions in realistic situations that are relevant to students' everyday lives. The questions

Table 1. Timeline of Activities – Year 1 (2007-08) of enVisionMATH RCT

Activities	August-September	October	November	December	January-February	March	April	May	June
Training and Program Implementation Begins	♦								
Assessments and Surveys Administered	♦							♦	♦
Site Observations	♦ (Site A-September)	♦ (Sites D, E, F, G, H)	♦ (Site B)	♦ (Site C)		♦ (Site F)	♦ (Sites A, B, C, H)	♦ (Sites D, E, G)	
Teacher Logs*		♦	♦	♦	♦	♦	♦	♦	

*Note that treatment and control teachers are completing monthly teacher logs that monitor instructional activities and the use of program and other resources.

assess students on their knowledge of mathematics as well as their ability to solve problems and communicate, and reason mathematically. The MAT8 aligns with the National Council of Teachers of Mathematics (NCTM) Standards 2000 and are reflective of what is taught in today's classrooms. The Math Computation, and Math Concepts and Problem-solving subtests were selected for administration. It took approximately 70 minutes to administer these subtests. The MAT8 provides raw scores, percentile ranges, and scale scores for each subtest. Split-half reliability coefficients[6] range from .84-.88 for each subtest, supporting the stability of the measures. The MAT8 has demonstrated validity (content, criterion, and construct-related validity). This information is described in detail in the publisher's technical manual.

- *Math Concepts and Problem-solving:* The Concepts and Problem-solving subtest (36 items at the 2nd grade level and 40 items at the 4th grade level) measure a student's facility for applying mathematics to many different kinds of problems and evaluating their results. Content covered includes numbers and operation, patterns and relationships, geometry, measurement, data and probability, and problem-solving skills. Estimation is also one of the important skills assessed in the Concepts and Problem-solving test.

- *Math Computation:* The 30 item computation test measures students ability to complete arithmetic operations. Depending on the level, this includes addition, subtraction, multiplication, and division of different kinds of numbers (whole numbers, decimals, fractions, and percents). Many of these items are presented in context so that the student must select an appropriate operation as well as perform the computation.

Group Mathematics Assessment and Diagnostic Evaluation (GMADE)- The GMADE, published by Pearson Assessment, is a norm-referenced, standards-based assessment of mathematical skills. According to the publisher, the GMADE is based on highly reliable scientific research and uses the standards set by the NCTM and is correlated to the new NCTM Focal Points. The Concepts and Communication subtest was selected for administration. This subtest uniquely addresses the language, vocabulary, and representations of mathematics. A symbol, word, or short phrase is presented with four choices. The four choices maybe pictures, symbols or numbers. The subtest contains a total of 28 items and takes approximately 15-20 minutes to complete. Reliability estimates range from .91 to .96. Raw scores, percentile rankings, and scale scores are available for this subtest.

Balanced Assessment of Mathematics (BAM)- Two versions of the BAM were used as part of this RCT--a published version for the 4th grade and one created for 2nd graders by PRES researchers. Both assessments are designed as performance assessments in order to provide students with an opportunity to show what they know and understand. Students are provided with a variety of multi-step tasks to measure mathematical performance.

- *The 4th grade BAM*, published by CTB and based on the work by the Harvard Group's Balanced Assessment in Mathematics Project, is designed to measure students performance against national and international standards. According to the publisher, the broader range and greater depth of the tasks enables this assessment to

[6] These reliability coefficients are based on the Kuder-Richardson Formula #20.

recognize, and thus encourage, students' achievements in meeting these higher standards for mathematical performance. The test contains 5 tasks and takes approximately 40 minutes to complete. Internal consistency was measured at $r_{posttest}$=.71 and $r_{pretest}$=.69, supporting the stability of the test items.

- *The 2nd grade BAM* was developed by PRES researchers. Items that measured a range of NCTM standards, involved various types of tasks, and were engaging were selected or adapted from two major sources. These included the Elementary Grades Assessment: Balanced Assessment for the Mathematics Curriculum (2006) and NCTM's Mathematics Assessment Sampler: PreKindergarten – Grade 2 (2006). The test contained 7 tasks and took approximately 40 minutes to complete. Internal consistency was measured at $r_{posttest}$=.68 and $r_{pretest}$=.69, supporting the stability of the test items. For more detailed information on the national standards and content tapped by this assessment, please contact PRES Associates.

In order to facilitate comparisons between the two tests, percent correct was the metric used in the present study. However, performance levels are also reported.

SURVEYS

Student Surveys. A student survey was developed to gather information on attitudes that may be affected by their math program. Specifically, the survey was developed to primarily measure:

- Attitudes about school *(e.g. I like school.)*
- Attitudes about math *(e.g. I like math.)*
- Perceived math ability *(e.g. I'm good at math)*
- Math effort and motivation (e.g., *I try hard in math class*)

The survey also included items on parental knowledge and support, teacher support, classroom experiences, and in the Spring survey, satisfaction with the math program. These scales were included in order to obtain measures of the impact of the enVisionMATH program on affective student outcomes and to measure potential variables that may serve as covariates as needed (e.g., parental support). While some items were created by PRES Associates, others were derived from scales with published reliability and validity[7]. Internal consistency of the scales measuring attitudinal constructs range from .49 to .90. High scores represent a very positive attitude or strong agreement (scales are from 1 to 5).

Teacher Surveys. A teacher survey was developed in order to obtain information on attitudes and reported practices that may be impacted by their math program. Specifically, the survey was developed to measure:

- Current and past classroom practices
- Math-related preparation and knowledge
- Confidence and comfort in teaching math
- Organizational factors/context
- Attitudes about student learning (inquiry versus traditional approaches), effective mathematics instruction, and their classes
- Attitudes about their curriculum

In addition, background and demographic information (e.g., years of experience,

[7] For the most part, student surveys for elementary children that are reliable and valid are difficult to find. In addition, they tend to be old. Therefore, a subset of items were selected from the entire survey and modified to be consistent with today's language. Survey information can be obtained from the following sources: Hogan, T. P. (1975). *Manual for Administering and Interpreting the Survey of School Attitudes*. New York: Hartcourt Brace; Johnson, O. G. (1976). *Tests and Measurements in Child Development: Handbook II.* San Francisco: Jossey-Bass; Marsh, H. (1990). The structure of academic self-concept: The Marsh-Shavelson model. *Journal of Educational Psychology, 82*, 623-636; Mulhern, F. & Rae, G. (1998). Development of a shortened form of the Fennema-Sherman Mathematics Attitudes. *Educational and Psychological Measurement, 58*, 295-306.

education, etc.) was also collected as potential covariates. Some items were obtained from existing scales, while others were developed for the study[8]. Internal consistency of the scales measuring attitudinal constructs range from .65 to .92. High scores represent a very positive attitude or strong agreement (scales are from 1 to 5).

Classroom Observation Forms. A classroom observation form was developed to guide observations. This form is largely based from existing protocols that have been used extensively and across the nation[9]. Modifications were made to reflect content and practices typical of elementary math classes, as well as to examine implementation of key components of enVisionMATH. Researchers conducting site visits and using classroom observation forms were trained extensively until a high level of agreement (.89 and above) was demonstrated among observers on the various quantitative and qualitative items.

Procedures

To ensure that all treatment teachers participating in the study had sufficient knowledge and skills to successfully implement enVisionMATH, teachers were given implementation guidelines and provided training prior to implementation. Such training is consistent with what could be expected in real-world settings. In addition, monitoring procedures (via monthly instructional logs completed by teachers and classroom observations and interviews) were developed to measure the extent to which treatment teachers were implementing a similar instructional model as outlined by the enVisionMATH program implementation guidelines. Of note is that while most of the imperative program materials, including Teacher's Editions and student books, were received before the start of the study, several teachers did not receive some of the program's ancillary components until several weeks after the start of the school year because they had not been released for distribution. However, all program materials were received as of the end of October, 2007.

The following section presents the procedures used to assist treatment teachers in implementing the enVisionMATH program, the monitoring procedures used by evaluators to determine treatment fidelity, methods used to obtain program feedback, and the test administration and scoring procedures employed.

TRAINING

The training model for the enVisionMATH RCT was carefully designed to provide treatment teachers with the necessary background to begin implementing the program right from the start. Additionally, all trainers assigned to train a participating study site attended a "train the trainer" meeting so that they were all clear on the information that needed to be conveyed to teachers during their trainings. This also helped to promote the consistency of trainings provided across all study sites. Teachers met with a Pearson professional trainer for approximately 5-6 hours prior to implementation of the program in their classes. During the initial training, the trainers clearly described the philosophy of the program and provided an overview of all

[8] Items in this survey were developed by PRES Associates and modified from the *Trends in International Mathematics and Science Study (TIMSS) 2003 Teacher Questionnaire Science Grade 8* (Washington, DC: National Center For Education Statistics) and the *2000 National Survey of Science and Mathematics Education Science Questionnaire* (Rockville, MD: Westat).

[9] The Classroom Observation Form was derived largely from the following protocols: Horizon Research's *Local Systematic Change Professional Development Classroom Observation Protocol*; Western Michigan University Science and Mathematics Program Improvement Center's *K-12 Mathematics Teaching Practices Observation Form;* and NSF-funded Collaborative for Excellence in Teacher Preparation's *Core Evaluation Classroom Observation Protocol.*

program components. Trainers also discussed which components were key (and required) versus those that were optional. Handouts (including the implementation guidelines) were also provided. These included materials lists, details on topics that should be clustered together, and suggestions on lesson flow. One particularly effective portion of the trainings occurred when trainers modeled a sample lesson for teachers to show them how the lesson should flow, language to use, and gave teachers a chance to see the program in action.

In addition to the initial, in-depth training, follow-up sessions were conducted at each site. The follow-up training sessions were designed as follows:

- Training Session Two (3 Hours), within 2 – 3 weeks of start of school year
- Training Session Three (3 hours), 8 - 10 weeks into school
- Training Session Four (4 hours), 12 - 14 weeks into school

The follow-up training sessions were somewhat less formal than the initial training and allowed opportunities for teachers to ask questions and receive feedback. In the majority of cases, during Training Session Two (first follow-up training) the trainers observed the treatment teachers using the enVisionMATH program during the first part of the day and conducted the training later in the day. It should be noted that while the training sessions were designed to take place as described above, some schools did not receive the full training regiment of 4 sessions. This was primarily due to difficulties in scheduling between busy teachers, schools, and trainers. However, in some cases, the school opted not to have a fourth training session as they felt that additional training was unnecessary at that point in the school year and teachers had a sufficient comfort level with the program. Table 2 shows training received by each site.

Table 2. Training Sessions by Site

	Training Session 1: Initial	Training Session 2: Follow-up	Training Session 3: Follow-up	Training Session 4: Follow-up
Site A	♦	♦	♦	
Site B	♦	♦	♦	
Site C	♦	♦	♦	♦
Site D	♦	♦		
Site E	♦	♦		
Site F	♦	♦	♦	
Site G	♦	♦	♦	♦
Site H	♦	♦		

Another item of note is that the focus of these trainings was not on general math professional development but rather on the vision of the enVisionMATH program and how the program can best be used to effectively help students make sense of mathematics.

IMPLEMENTATION GUIDELINES

Teachers were provided with detailed implementation guidelines to ensure that all treatment teachers participating in the study had a clear understanding of the key program components of the enVisionMATH program. Pearson staff and program author, Randy Charles, identified key components of the 2009 *enVisionMATH* program which formed the basis of the implementation guidelines. The guidelines offer detailed direction on how the program should be used in the classroom as well as what part of the program are considered key (and required), versus what program elements are considered optional. The key components of the program include:

- Review of the math/lesson background sections (Quick and Easy Lesson Overview, Mathematics Background section in TE)

- Problem of the Day
- Daily Spiral Review
- Interactive Learning (IL) activity
- New vocabulary
- Visual Learning Band
- Set the Purpose
- Guided Practice
- Quick Check
- Assignment of independent practice
- Differentiated Practice/Instruction (e.g., Leveled homework)
- Close in TE
- Error Intervention/ Reteaching/MDIS

For a full description of these key components, please see Appendix E.

PROGRAM MONITORING

Teacher Logs. Online teacher logs were used so that program implementation could be monitored on a real-time basis and to identify any issues or local events that may influence study results. Teachers were instructed to complete these on a monthly basis from October through May. Both treatment and control teachers completed teacher logs, with slightly different versions for each. The primary purpose of the teacher logs for treatment teachers was to monitor program implementation and fidelity. The reason researchers also collected monthly logs from control teachers was so instructional activities and content covered could be monitored, and also to monitor the extent to which any contamination may have occurred. Such background information provided researchers with a detailed data source on what was occurring in treatment and control classrooms in terms of math instruction and allowed researchers to identify areas of overlap in terms of content taught and activities. The extent to which there are similarities and differences between classrooms can have an impact on observed differences between treatment and control classes and effect sizes. Thus, it is important to take these factors into consideration when interpreting study results. Information obtained via these logs included changes in their student roster, typical classroom activities, use of other resources and exercises (including homework and independent practice), and for treatment teachers, use of key enVisionMATH program components.

Results showed that teachers had, on average, a 88% completion rate. The ranges were 0% to 100%[10]. Teachers were contacted after failure to complete teacher logs each month. In cases of noncompliance, the school liaison was asked to confer with the teacher to see if there was anything that could be done to assist the teacher in completing the logs. Furthermore, for those teachers that did not have high completion rates, a more extensive implementation checklist and interview was completed during the Spring site visit to ensure that information on implementation, instructional practices, and classroom activities was available on that teacher.

Classroom Observation. Classroom observations were conducted for all treatment and control teachers during the Fall (October-December, 2007) and the Spring (March-April, 2008). The purpose of these observations was to better understand the instructional approaches and materials used by teachers with their students and to identify differences and similarities between teachers who were randomly assigned to treatment and control conditions[11]. Specifically, observations focused on how

[10] Calculation based on 8 months in which teachers were asked to report on their activities. One teacher refused to complete logs. However, information was collected from this teacher via more extensive interviews conducted throughout the school year. Note that this high completion rate was obtained, in part, due to the added monetary incentive given to teachers for completion of data collection activities.

[11] It should be noted that random assignment alone does not ensure initial group equivalence, especially with small sample sizes (e.g., less than 1000,; Lipsey, 1990).

classroom activities were structured, what and how materials were used, and characteristics of the class including student engagement, classroom environment and culture, and teacher-student interactions. In addition, teachers were interviewed after the observations to obtain more specific information on the representativeness of the lesson, resources used, ability levels of the students, assessment practices, pacing, independent practices, and test preparation strategies. The observations also allowed researchers to examine the extent to which teacher level differences could have influenced study results and to examine the threat of possible contamination between treatment and control teachers.

TEST/SURVEY ADMINISTRATION AND SCORING

All assessments were administered by teachers during the Fall (August through September, 2007) and Spring (May through June, 2008)[12]. The test publisher's standard testing procedure was distributed to all teachers prior to testing. Teachers were instructed to contact PRES Associates if they needed further guidance. The MAT8 and GMADE tests were scored by PRES researchers following the standardized scoring procedures (including raw score conversions) as outlined in the publisher's technical/scoring manual. Scorers were unaware of group assignment.

The BAM tests were scored by one research assistant following extensive training and calibration with a senior researcher experienced in scoring rubric-scored assessments. Only until the research assistant and senior researcher showed 95% consistency in scoring for the sample tests did the research assistant proceed to score official tests collected as part of the RCT. The scorer was blind to group assignment.

Student and teacher surveys were completed during the Fall (September through October, 2007) and Spring (May through June, 2008).

Curricula

As part of the initial site selection criteria, researchers tried, to the extent possible, to select schools which used a control program that differed pedagogically from the intervention under study. Indeed, prior to approving a site for participation, the control curricula were evaluated to determine if the program was sufficiently distinct. For the enVisionMATH RCT, participating schools are using a variety of curricula from other publishers. However, it is also important to note that teachers are all teaching similar math concepts and due to state and national guidelines, are covering similar content (e.g., addition, fractions, division, etc.); thus there are similarities in content covered between treatment and control programs. The focus of this study is to examine the effects of an entire curriculum (enVisionMATH) and as such, it must be compared to other curricula that teach the same content area.

2009 ENVISIONMATH PROGRAM

The 2009 enVisionMATH program is a research-based instructional model designed to make mathematics more accessible to a wide range of students. Through interactive learning and problem-based activities, students are able to build their own understanding of concepts and skills before the formal representation of ideas occurs. Visual representations drive concept and skill development and each lesson contains a student "visual learning band" which

[12] Administration dates depended on the school's start and end date. Teachers within each school followed a similar testing schedule. Generally, administration occurred within 1 month after the school year commenced (pretest) and within 1 month prior to the end of the school year (posttest).

incorporates a dynamic presentation of the objective and essential understanding of the lesson. Timely, frequent assessments assist teachers in individualizing instruction, which is supported by the large range of differentiated instructional resources provided to teachers. Technology alternatives allowed the print version to come alive through motion and sound. Teacher explanations and Center Activities reinforce, deepen and extend learning.

The enVisionMATH program is organized into 20 individual content topics, rather than longer, broader chapters. Each topic contains four to nine lessons and develops one or a few related content standards in depth. enVisionMATH was developed so that all of the lessons in the program can be taught prior to the end of year state/district testing.

To accomplish the goals of the enVisionMATH program, resources were carefully designed to meet the needs of all students. Some of the ancillary materials included with the program and used by participating teachers include:

- Interactive Homework Workbook
- Interactive Math Series Big Book (K – 2)
- Math Diagnosis and Intervention System
- Individual Student Manipulative Kits
- Teacher Overhead Manipulative Kits
- Center Activities Kits
- Visual Learning Bridge Transparencies
- MathStart readers by Stuart Murphy (K – 2)
- World Scape readers (3 – 6)
- ExamView
- ETools

It should be noted that several built-in components of the enVisionMATH program, such as the Math Diagnosis and Intervention System (MDIS), are designed to aide teachers in reaching all learners. The MDIS and other program tools help teachers provide individual instruction and remediation to students below level. More detailed information on how this program is being implemented by teachers as part of this study is provided in the section "Fidelity of Program Implementation" and in the Implementation Guidelines in Appendix E.

CONTROL CURRICULA

There were eight distinct control programs used by teachers participating in the enVisionMATH study, with two schools using state versions of the same curriculum and one school using different programs in the second and fourth grade. The components of the control programs were varied somewhat, particularly given that many control teachers supplemented their core programs with other materials and did not exclusively use the core math programs provided (this was particularly true of schools F and G). Table 3 provides a summary of the different types of control programs used at each of the sites.

With the exception of program 1 (school A), program 6 (school G) and program 7 (used in the second grade at school H), the control curricula used at participating schools were similar to one another in that they were all more traditionally-based math programs from well-known publishers. Each of these programs were very similar to one another in terms of overall approach and design; they all emphasized similar pedagogical approaches and contained features that were comparable. Commonalities between programs 2, 3, 4, 5, and 8 included:

- Lessons begin with some type of warm-up or activity to engage students, followed by lesson instruction and then practice.

- Provide leveled practice and suggestions for differentiated instruction to meet the needs of students of various ability levels.
- Includes hands-on activities to stimulate students' interest in math.
- Problem-solving is emphasized and incorporated into most lessons.
- Incorporates built-in program assessment.

Though these control curricula shared similarities, there were a few notable differences between control programs 2, 3, 4, 5, and 8. As mentioned above, the same general lesson pattern was followed for these curricula; however, program 5 also included a lesson section specifically for differentiation, and programs 3 and 8 emphasized checking understanding and reteaching prior to independent practice. Additionally, program 5 did not provide manipulatives as part of the program, so while there were some hands-on activities incorporated into lessons, these were included to a lesser extent.

In contrast to the more traditional programs, program 1 at school A was very focused on developing students' math skills through activities. The program employed a constructivist approach and instruction revolved around hands-on experiences and real-world applications of math. The second grade program at school H relied on a basal program (though not chapter-based) that focused on providing repetition, cumulative review and distributed practice. While investigative elements were included in the program, much more lesson time is devoted to independent practice. The other unique control curriculum was program 6. The type of materials used for program 6 at school G varied widely and included a variety of math programs, teacher created materials, or other resources collected over the years. One teacher used a much more traditional approach, while the other used a constructivist approach to math instruction that emphasized students discovering math on their own through investigative activities.

In general, the content covered in the control classes at all sites was comparable. The scope and sequence of all programs covered similar topics. While entire chapters may not have been dedicated to the same content across all programs, the majority of math topics were covered to some extent by each program. Therefore, while the sequence of the content varied from program to program, similar content was taught. Additionally, it should be noted that the majority of control teachers had been using their programs for over 3 school years.

Table 3. Control Curricula by Site*

	Site A	Site B	Site C	Site D	Site E	Site F	Site G	Site H
Prog. 1 Investigative	2nd & 4th							
Prog. 2 Basal		2nd & 4th						
Prog. 3 Basal			2nd & 4th					
Prog. 4 Basal				2nd & 4th	2nd & 4th			
Prog. 5 Basal						2nd & 4th		
Prog. 6 Mix							2nd & 4th	
Prog. 7 Basal								2nd only
Prog. 8 Basal								4th only

*To ensure confidentiality, the actual names of the control curricula are omitted. Appendix C provides more detailed information on the characteristics of control curricula, content covered and materials used by teachers.

Indeed, 37% of teachers reported they had been using their math program for 3-5 years and 33.3% had been using their program for more than five years.

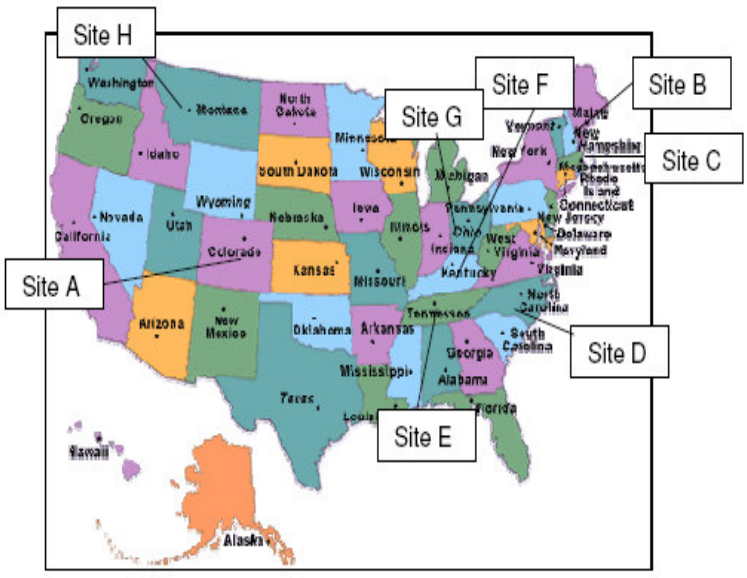
Site Selection Criteria

Sites were selected using the following criteria:

- Diverse geographical areas;
- Schools had to contain multiple teachers at the 2nd through 5th grades;
- Historically low student mobility rates (less than 20%);
- Interest in using enVisionMATH; and
- Willingness/commitment to fully participate in all aspects of the study

Other major criteria included: 1) that there be no other major math initiative(s) at the school; and 2) the typical math curriculum employed by the school fell under the "comparison" programs which provided somewhat of a contrast to the enVisionMATH program.

Figure 1. 2007-08 enVisionMATH RCT Study Sites



Sample Description

Site Characteristics

A total of 8 elementary schools participated in the first year of the study. Figure 1 displays the geographical location of each of the participating sites[13]. As shown, sites are geographically dispersed across the United States.

A detailed case study of sites is provided in Appendix B. Table 4 on the following page shows characteristics of each of the participating sites. As shown, schools participating in the RCT are predominantly White, with approximately 13% of students being economically disadvantaged. However, there is also some variation between sites; e.g., whereas sites B and C have less than 10% of economically disadvantaged students, sites D, E and H have over 20% students who are designated as such. This variation in sites allows researchers to examine the effects of enVisionMATH on subgroups of students and types of schools.

Teachers reported that overall the classes included in the study contained a broad-range of abilities, with some variance between classes. Furthermore, monitoring of sites showed no evidence of a local history event or major disruption.

[13] Due to confidentiality agreements with each of the participating sites, the exact location and name of the schools have been omitted.

Table 4. Student Demographic Characteristics by School

School	School Size	Teachers By Group	Total Students by Group and Grade level	Ethnic Breakdown	% Special Education	% Limited English Proficiency	% Free/ Reduced Lunch	Gender Breakdown
Site A ***Colorado*** **Grades K-6**	417	6 4-TX 2-CT	129 36-TX 2nd grade 19-CT 2nd grade 50-TX 4th grade 24-CT 4th grade	71.8% White 19.4% Hispanic 5.6% African Am. 3.2% Native Am.	10.4%	1.6%	NR	53.2% Male 46.8% Female
Site B ***New Hampshire*** **Grades Pre-K-4**	569	12 6-TX 6-CT	268 63-TX 2nd grade 64-CT 2nd grade 68-TX 4th grade 73-CT 4th grade	94.8% White 2.2% Hispanic .7% African Am. 2.2% Asian/Native Am./Other	5.2%	.4%	10.1%	52.6% Male 47.4% Female
Site C ***Massachusetts*** **Grades PreK-5**	733	15 6-TX 9-CT	333 66-TX 2nd grade 89-CT 2nd grade 67-TX 4th grade 111-CT 4th grade	94.6% White 1.5% Hispanic 1.8% African Am. 1.5% Asian	17.1%	1.2%	2.1%	49.2% Male 50.8% Female
Site D ***North Carolina*** **Grades PreK-5**	544	6 3-TX 3-CT	87 26-TX 2nd grade 17-CT 2nd grade 21-TX 4th grade 23-CT 4th grade	80.5% White 19.5% Hispanic	25.3%	NR	NR	47.1% Male 52.9% Female
Site E ***Tennessee*** **Grades PreK-5**	249	4 2-TX 2-CT	75 17-TX 2nd grade 16-CT 2nd grade 20-TX 4th grade 22-CT 4th grade	97.3% White 2.7% Hispanic	17.2%	2.7%	21.3%	58.7% Male 41.3% Female

TX=treatment, CT=control, NR=not reported

*Table 4 continued on following page.

Table 4 Continued. Student Demographic Characteristics by School

School	School Size	Classes By Group	Total Students by Group and Grade level	Ethnic Breakdown	% Special Education	% Limited English Proficiency	% Free/ Reduced Lunch	Gender Breakdown
Site F ***Kentucky*** **Grades PreK-5**	514	8 4-TX 4-CT	174 40-TX 2nd grade 40-CT 2nd grade 48-TX 4th grade 46-CT 4th grade	96.0% White 1.7% Hispanic 1.7% African Am. 0.6% Asian	9.2%	NR	NR	53.4% Male 46.6% Female
Site G ***Ohio*** **Grades K-6**	270	4 2-TX 2-CT	68 17-TX 2nd grade 16-CT 2nd grade 20-TX 4th grade 21-CT 4th grade	100.00% White	5.9%	1.5%	25.0%	55.9% Male 44.1% Female
Site H ***Montana*** **Grades K-6**	240	4 2-TX 2-CT	63 15-TX 2nd grade 16-CT 2nd grade 15-TX 4th grade 17-CT 4th grade	94.6% White 5.4% Native Am.	3.6%	0.0%	69.6%	64.3% Male 35.7% Female
Sample Characteristics (Across all sites)		59 29-TX 30-CT	1197 279-TX 2nd grade 273-CT 2nd grade 310-TX 4th grade 335-CT 4th grade	91.9% White 4.8% Hispanic 1.5% African Am. 1.8% Asian/Native American/Other	11.8%	1.1%	13.2%	52.6%Male 47.4%Fem.
National Population[14]				**White-61.2%** **Hispanic-16.3%** **Black-17.2%** **Asian-4.1%** **Other-1.2%**	**13.6%**	**10.6%**	**40.9%**	**Male-50.2%** **Female-49.8%**

TX=treatment, CT=control, NR=not reported

[14] Data was obtained from SchoolDataDirect.com, an online service of the Council of Chief State School Officers' State Education Data Center. Figures represent distributions across all grade levels and reported for 2006, (NCES 2002-130).

Student Characteristics

The final sample consisted of 1197 students (608 control; 589 treatment) in 59 classes (30 control and 29 treatment). The 2007-08 participating students were in the 2nd (46.1%) or 4th grade (53.9%). Table 5 presents the demographic distribution among study participants. Note that only students who remained in the study throughout the year are included in this table and in the final analyses. As previously noted, the sample includes a higher proportion of White students as compared to those found nationwide.

Preliminary analyses[16] were performed to examine whether baseline differences existed as a function of student demographics. Chi-square analyses on the demographic characteristics noted in Table 5 showed that only ethnicity (White, minority) was significantly associated with group,

Table 5. Student Demographics Distributions*

Characteristics		Control (n=608)		Treatment (n=589)		Total (n=1197)		National[15]
		Count	Percent	Count	Percent	Count	Percent	Percent
Gender *($\chi^2(1)=.49$, $p=.49$)*	Male	310	51.6%	312	53.6%	622	52.6%	50.9%
	Female	291	48.4%	270	46.4%	561	47.4%	49.1%
Ethnicity *($\chi^2(3)=12.19$, **$p=.007$**)*	White	570	94.5%	519	89.2%	1089	91.9%	55.9%
	Hispanic	18	3.0%	39	6.7%	57	4.8%	20.5%
	African Am.	6	1.0%	12	2.1%	18	1.5%	16.9%
	Asian/Native American/Other	9	1.5%	12	2.1%	21	1.8%	5.7%
Grade *($\chi^2(1)=.73$, $p=.39$)*	2nd	273	44.9%	279	47.4%	552	46.1%	--
	4th	335	55.1%	310	52.6%	645	53.9%	--
Subpopulations								
($\chi^2(1)=.003$, $p=.96$)	Special Education	71	11.8%	67	11.9%	138	11.8%	13.6%
($\chi^2(1)=.34$, $p=.56$)	Free/Reduced Lunch Status	55	12.6%	51	14.0%	106	13.2%	40.9%
($\chi^2(1)=1.89$, $p=.17$)	Limited English Proficiency	3	.6%	7	1.6%	10	1.1%	10.6%
($\chi^2(2)=1.99$, $p=.37$)	Low Math Level	136	22.8%	132	22.7%	268	22.7%	--
	Average Math Level	169	28.3%	145	24.9%	314	26.6%	--
	High Math Level	292	48.9%	305	52.4%	597	50.6%	--

* Counts (and percents) do not include missing information. Ability level was determined by using pretest percentile rankings. Students who were at or below the 33rd percentile were classified at a low math level, students who were at or above the 66th percentile were classified as high, and the remaining students were classified as average. Results showed that 70 students missed more than ¼ of the school year (41 treatment and 29 control students); this includes students enrolling in school later in the school year, leaving the school, or due to absenteeism There was no relationship between group and attendance, $\chi^2(1)=2.32$, $p=.13$. These students are excluded from the above table and subsequent analyses.

[15] Data was obtained from SchoolDataDirect.com, an online service of the Council of Chief State School Officers' State Education Data Center. Figures represent distributions across all grade levels and reported for 2006, (NCES 2002-130).

[16] All details regarding analyses on baseline differences and attrition analyses are provided in the accompanying Technical Report. Main findings are drawn from the Technical Report and presented herein.

$p<.05$[17]. Specifically, results showed that there were more White students in the control group than the treatment group. Therefore, this variable was used as a covariate in subsequent analyses of program effects.

Differences on other student characteristics were also examined. Results showed no significant differences between treatment and control students in terms of perceived parental involvement and support, liking for school, and school effort and motivation, $p>.05$.

In terms of ability levels, the majority of teachers reported that their students had a broad-range of ability levels, but on average, they were typical of other 2nd and 4th graders. The vast majority of classes were taught at grade level (as opposed to advanced (gifted) or remedial math classes). Examination of the pretest scores showed that in general, participating students were slightly above average as measured by the MAT8's Concepts and Problem-Solving and Computation subtests, with 2nd graders scoring at the 57th percentile and 4th graders scoring at the 60th percentile. On the GMADE, which measures math vocabulary skills, students performed below average (35th and 38th percentiles for 2nd and 4th graders respectively). On the BAM, which measures math problem-solving and communication, students showed a "Beginning Level of Math Understanding."

In addition, pre-test differences on the assessment measures were also examined, see Table 6. The student level analyses revealed that control students performed significantly better on the MAT8 Math Computation at pretesting compared to treatment students, $p<.05$. There were no differences on the remaining tests. It should be noted that examination of baseline levels via multilevel models revealed no significant differences between groups on all measures.

Table 6. Sample Size, Means, Standard Deviations, and *t*-test (Student Level) Results for Assessments at Pre-testing

Pretest*	Group	N	Mean	Std. Dev.	*t*	Sig. Level
MAT 8 Concepts and Problem-solving Scale Score	Control	581	609.40	51.49	.49	.62
	enV	574	607.96	47.99		
MAT 8 Computation Scale Score	Control	579	585.49	65.49	2.62	.01
	enV	572	575.83	59.30		
GMADE Scale Score	Control	589	94.08	10.93	1.57	.12
	enV	570	93.06	11.19		
BAM Percent Correct	Control	581	54.98	17.82	.49	.62
	enV	573	54.49	16.29		

[17] "Significant" means that we can be 95% or more confident that the observed differences are real. If the significance level is less than or equal to .05, then the differences are considered statistically significant. If this value is greater than .05, this means that any observed differences are not statistically significant and may be interpreted as inconclusive. However, at times this may be referred to as "marginally significant." In this case, the criterion is more liberal and means that we can be 90% or more confident that the observed differences are real. Throughout this report, significant differences are presented in bold.

Teacher Characteristics

There were 56 teachers (28 control and 28 treatment) who participated in the RCT. A total of 29 teachers taught the 2nd grade (15 treatment and 14 control) and a total of 27 teachers taught the 4th grade (13 treatment and 14 control). Approximately 88% of teachers are female and 12% are male. Teachers are primarily Caucasian (92%).

In regards to educational background, 28% of teachers have a Bachelor's and 72% have a Master's in Education, Leadership, Curriculum/Instruction, Reading/Language Arts or Special Education. With the exception of one teacher, all teachers are state certified to teach elementary education.

Teacher experience ranged from 0 to 32 years. Second grade teachers had taught from 0 to 24 years, with the average number of years taught being 5. Fourth grade teachers had taught from 1 to 37 years, with the average number of years taught being 8.5.

Preliminary analyses were run to examine whether the treatment and control teachers were comparable on key teacher level characteristics[18]. Results showed no significant baseline differences among teachers in terms of knowledge of NCTM standards and focal points, preparation to teach math via "best practices" strategies, hours of professional development received over the last three years, number of formal courses taken in mathematics, and teaching experience in the grade they are currently teaching. There were also no differences on affective measures such as perceptions of control over teaching, confidence in teaching math, beliefs about inquiry-based instructional practices, beliefs about traditional math instructional practices, perceived usefulness of mathematics, perceived teacher collaboration and support, perceived parental/administrator support, and access to ancillary resources to teach math. However, differences were observed. Specifically, control teachers perceived greater discomfort in teaching math than enVisionMATH teachers as measured by the pre-survey, a higher percentage of control teachers (85%) had earned a MA degree as compared to treatment teachers (58%), and control teachers had greater overall teaching experience, with an average of 16 years as compared to 10 years by treatment teachers.

Classroom environment and implementation of various typical activities that occur in elementary math classrooms were also analyzed based on information collected from the classroom observations, logs, and teacher surveys. Results showed no significant differences between treatment and control teachers in terms of implementation of math activities, student activities, teacher use of technological resources, assessment practices, homework practices (frequency, completion rate), teacher-student interactions, class culture, and class engagement.

In summary, although some differences were found between treatment and control teachers and students, randomization was reasonably successful in producing equivalent treatment and control groups in terms of student and teacher characteristics. Nevertheless, to enhance the analyses' power to detect effects and further equate groups, covariates were included in multilevel models examining program effects.

[18] Note that these analyses at the teacher level have an inherent lack of power to detect differences due to the small sample size (n=56). Detailed statistics are provided in the accompanying Technical Report.

Instructional Content, Coverage and Practices

As a result of district and state scope and sequence guidelines detailing what math content needed to be covered, treatment and control teachers within schools generally taught a similar amount of content. Comparison on the percent of chapters/topics completed during the school year, relative to the minimum that should have been completed (per the texts used), showed that while on average both 2nd grade treatment and control teachers taught a similar amount of content, 4th grade control classrooms tended to cover more of their math program (95%) than enVisionMATH classrooms (78%). This is likely due to the pacing issues that the treatment classrooms experienced early on as these teachers and their students were becoming accustomed to the program.

There were also some differences in the specific content covered between treatment and control teachers across schools. For example, 2nd grade control classrooms covered more multiplication than treatment classes. However, enVisionMATH 2nd grade classes tended to place more emphasis on understanding addition and subtraction, mental addition, and mental subtraction as compared to 2nd grade control classrooms. At the 4th grade, the enVisionMATH program included a chapter on numeration that was not included in many control programs. Furthermore, the 4th grade control program used at site D included a chapter devoted to dividing by 2 digit numbers and a chapter on fractions and decimals in circle graphs. While other programs covered this content, entire chapters were not devoted to this content within the remaining programs.

These were the only notable differences observed across schools in terms of content taught. However, for the most part, students were taught similar content. Appendix C contains a crosswalk between enVisionMATH content and the control programs' content. As is clearly evident, there exists a close alignment. This is largely due to the educational community's demand on publishers to include content that is aligned to national and state standards.

In addition to similar content being taught across treatment and control classes, the way that teachers structured their lessons was also very comparable. This is because most teachers, whether treatment or control, tend to use common approaches including warm-up and review, teaching new information and then application through independent practice. Indeed, treatment and control teachers showed no significant differences in the amount of "effective[19]" math pedagogical practices, teacher-student interactions, class culture, class engagement, and overall class climate. Common pedagogical practices (i.e., how lessons are delivered) employed by teachers included:

- Help students develop communication (reading/ writing) skills relevant to math.
- Provide intervention when students need it and individualize instruction to the needs/developmental levels of the students (e.g. for below average, average and advanced student).
- Help students understand previously taught concepts through review.
- Formally assess students' understanding via tests and/or quizzes.
- Make connections to the real-world.
- Require students to show their work and explain their math solutions.

[19] The teacher survey asked teachers to report on the extent to which they engaged in various math and general pedagogical practices, including those that have been shown to be effective in the research literature.

- Encourage students to learn and use math vocabulary.
- Encourage students to use different ways to arrive at math solutions.
- Make connections to previously taught concepts.

However, there were also some differences noted between treatment and control teachers that can be attributed to their respective math programs. For example, treatment teachers reported engaging in the following activities to a greater extent over the school year:

- Math-related hands-on activity (e.g., an investigation or interactive learning activity)
- Read a story that teaches a math concept or includes math-related ideas
- Use mathematical concepts to solve real-world problems.
- Engage in mental math (e.g. estimation, thinking about the solution, etc.)

Additionally, control teachers noted that they were more likely to practice basic math computations/algorithms and help students use worksheets, flashcards or other tools to practice memorization of math facts than treatment teachers. These variations in program-specific teaching practices are to be expected when teachers are using and adhering to distinct curricula. Indeed, as previously noted, the enVisionMATH program is unique in many respects as it tries to blend the best of constructivist and traditional math approaches. As such, comparisons between enVisionMATH and other programs are likely to yield differences in pedagogical practices. Indeed, this is examined in greater detail on page 39.

FIDELITY OF IMPLEMENTATION

Available information from observations, monthly teacher logs and teacher surveys were triangulated in order to determine the level of implementation fidelity among enVisionMATH teachers. Three levels of implementation (low, moderate, and high) were assigned for treatment teachers' implementation of key program components as outlined in the implementation guidelines and the percent of topics completed. This information was then averaged to determine an overall implementation level[20] of the enVisionMATH program in terms of breadth and depth. As shown in Table 6, high fidelity of implementation was evident among the majority of treatment teachers. That is, teachers tended to use the program as described in the implementation guidelines and also covered a number of topics. That said, there were some key program components that were used with less frequency by some teachers. This included the Close section of the Teacher's Edition, assigning Leveled Homework to students, and reviewing the Mathematics Background section of their Teacher's Edition. It is also of note that 45% of teachers were unable to complete most of the 20 topics during the school year. This is evidence of early pacing issues teachers experienced as they were first becoming familiar with the new enVisionMATH program. It must be emphasized that lack of program completion will diminish the effects sizes observed as a result of enVisionMATH. That said, as a two-year study, it is expected that 3rd and 5th grade treatment teachers who were given the option of using enVisionMATH in 2007-08, will not experience such pacing issues since they are already familiar with the program.

[20] Researchers discussed and reviewed the available data to assign teachers to a level of implementation when variation between completion of key program components and topics completion existed.

This can be expected to contribute positively to the impacts of enVisionMATH during the second year of the evaluation.

Table 7. Level of enVisionMATH Implementation

Level of enVisionMATH Implementation	Completion of Key Program Components	Mean Percentage (and number) of enVisionMATH Topics Completed
High = 20 teachers	85% of goals met (i.e., consistent completion of 11 or more out of 13 enV components)= **24 teachers**	90% or higher (18 or more) = **16 teachers**
Moderate = 5 teachers	69-80% of goals met; (i.e., consistent completion of 9 to 10 out of 13 enV components) = **3 teachers**	70%-85% (14-17) = **9 teachers**
Low = 4 teachers	Less than 69% of goals met = **2 teachers**	65% or less (<=13) = **4 teachers**

- *Note that 86% of teachers implemented enVisionMATH with a moderate to high level of fidelity (with high being the norm).*

No evidence of contamination was observed among control teachers at any of the participating sites. That is, control teachers did not use any components of the enVisionMATH program with their students. However, two teachers at school C who were initially assigned as treatment teachers were dropped from the study due to extremely low implementation of the enVisionMATH program. These teachers supplemented enVisionMATH so heavily that it no longer resembled the enVisionMATH program. As a result, the teachers were excluded from the study late in the Fall[21].

It should be noted that the potential for contamination was given careful consideration when determining the level of random assignment. Through years of research experience, PRES researchers have found that the benefits of random assignment at the teacher level (hence, controlling for school level factors) *with* careful monitoring of possible contamination, outweighs the risk of contamination. Procedures used to eliminate the threat of contamination included an in-depth study orientation with both treatment and control teachers, site visits made to both treatment and control classrooms to observe what was occurring in classrooms, and monthly teacher logs that monitored practices and materials used.

Overall, treatment teachers implemented the key enVisionMATH program components with a high degree of fidelity. However, 45% of teachers were unable to complete all 20 topics during the 2007-08 school year due to initial pacing issues associated with using a new program. This is expected to improve during the upcoming school year as 3rd and 5th grade teachers who have been using the enVisionMATH program for more than one year will constitute the majority of the treatment teacher sample in year 2.

Results

This section is organized by the key evaluation questions and provides a summary of major findings first, followed by a more detailed account of the results. The findings described in this report provide a summary of overall conclusions that can be derived from the extensive analyses conducted. However, detailed descriptions of the statistical analyses performed along with detailed statistical results are provided in the accompanying *Technical Report*.

[21] Note that while researchers contacted the teachers and attempted to persuade them to use more of the enVisionMATH program, the teachers did not wish to do so.

Summary of Results

Does math ability improve over the course of participating in enVisionMATH? Does this vary by different types of students and levels of implementation?

Students using enVisionMATH significantly improved over the course of the school year in the areas of math concepts and problem-solving, math computation, math vocabulary, and communication in math. Specifically, results showed that enVisionMATH students demonstrated significant percentile gains of:

- 19 points on concepts and problem-solving,
- 33 points on computation,
- 15 points on math vocabulary, and
- 36 points on math communication.

Another way to look at norm-referenced assessment results is to look at the percentile rankings of students relative to a national sample. It is a general rule of thumb that if a student makes a year's growth for a year of instruction, then the percentile rank will remain the same. enVisionMATH students had higher percentile rankings at post-testing than pre-testing on the MAT8 Computational subtest (54^{th} to 58^{th} percentile) and the GMADE (36^{th} to 45^{th} percentile). This means that enVisionMATH students grew more than would be expected over the course of a typical academic year as measured by the math computation and vocabulary tests.

Furthermore, the enVisionMATH program worked just as well with 2nd and 4th graders, females and males, White and non-White students, special education and non-special education students, students receiving free/reduced lunch and those not receiving this aid, and students at various math levels. Although a greater rate of improvement was demonstrated for certain subgroups of students, enVisionMATH students in all subpopulations showed significant learning gains on all assessment measures.

Since there was some variation observed in overall implementation of the enVisionMATH program among treatment teachers, analyses were performed to examine if this affected student performance. Results showed that there was a significant relationship between overall enVisionMATH implementation levels and improvement in math performance. Specifically, preliminary analyses showed that students whose teachers implemented the major components of enVisionMATH with high fidelity showed greater improvement as compared to students of teachers who implemented enVisionMATH with low fidelity and did not use the major program components on a regular basis.

How does mathematics performance differ between students who use enVisionMATH as compared to students who do not use this program? Do effects on math performance differ across types of students or settings?

Results showed positive effects of the enVisionMATH program. Elementary students who used enVisionMATH in 2007-08 showed greater gains in math computation, math vocabulary, and math problem-solving and communication as compared to students who used other math programs. In addition, while effects were small (ranged from .20 to .24), they are both significant and meaningful – especially given the short duration of the study and the applied settings in which the study was undertaken. Given that this study will continue into a second year (2008-2009

school year), larger effects could be expected if such trends continue and students are exposed to enVisionMATH for even longer periods of time.

Analysis of subgroup differences also showed significant effects. Notably, the following types of enVisionMATH students showed significantly greater gains in math performance as compared to control students in these subgroups:

- 4th graders
- Minorities
- Females
- High math ability students

Moreover, of the 16 significant subgroup differences observed, 14 were in favor of the enVisionMATH program (i.e., enVisionMATH students outperformed control students).

In addition, the positive effects obtained on the enVisionMATH program were observed across a number of different schools who used a variety of types of control programs. Specifically, enVisionMATH students performed significantly better than control students who used programs that were purely investigative and inquiry-based as well as students who used more traditional basal math programs. This consistency in findings across different curricula, schools, and measures lends credence to the conclusion that enVisionMATH positively impacts student math knowledge and skills.

Does participation in enVisionMATH result in other positive student outcomes (e.g., positive attitudes towards math, etc.)?

While the main focus of the enVisionMATH program is to improve upon important math skills and understanding, other measures were included to explore if enVisionMATH was associated with positive impacts on student and teacher attitudes, and classroom practices. Results showed that enVisionMATH students enjoyed math more and perceived greater teacher support as compared to control students. enVisionMATH and control students had similar positive attitudes in various other areas (e.g., importance of math, attitudes about math success, and so forth).

Results also showed that enVisionMATH had positive effects on teacher attitudes and instructional practices. In particular, enVisionMATH teachers noted that they were more prepared to carry out various mathematics activities, and in fact, tended to engage in a greater variety of mathematics activities and strategies as compared to control teachers. enVisionMATH teachers also reported greater confidence in their ability to teach math as compared to control teachers and showed a significant increase in their perceived knowledge of NCTM standards and focal points. In addition, following one year of enVisionMATH use, teachers tended to be more likely to incorporate inquiry-based strategies into their math instruction. That said, the overall pedagogical approach employed by enVisionMATH teachers tended to blend inquiry-based and traditional approaches. This is consistent with the overall enVisionMATH philosophy in that it attempts to blend both types of pedagogies.

What did users of enVisionMATH think about the program?

The enVisionMATH program was highly regarded by the vast majority of teachers. A full 92% of enVisionMATH teachers surveyed indicated that the program was an effective tool for mathematics instruction. In addition, 89% of enVisionMATH teachers indicated that they would recommend their

math program in comparison to only 39% of control teachers. Comparisons also revealed that enVisionMATH teachers rated their math programs' resources (e.g., professional development embedded within program, reading/writing in math activities, technology resources, review materials, etc) as more useful compared to control teachers. Teachers noted a variety of specific program components when asked to identify the three things they liked best about the enVisionMATH program. However, a few items emerged as favorites, including:

- The amount of story problems and the "thinking about math" involved in the program.
- The Daily Spiral Review
- The visual representation of concepts throughout the program
- The manipulatives provided and hands-on aspects of the program
- The amount of writing provided in the program
- The design of the pouches and Teacher's Edition
- The technology and online resources provided

Students who used enVisionMATH also enjoyed the program. Significant differences were observed in that enVisionMATH students rated their math program higher than control students; for example, 73% of enVisionMATH students noted that they liked the program used in math class as compared to 62.8% of control students.

Detailed Results

Does math ability improve over the course of participating in the enVisionMATH program?

In order to examine the extent to which enVisionMATH students experienced learning gains in mathematics, paired sample t-tests and multilevel modeling were performed[22]. Results showed significant growth in math knowledge and skills among enVisionMATH students (across both grade levels) as measured by all assessments, $p<.05$, see Figures 2 and 3. That is, students who were taught with enVisionMATH exhibited significant learning gains from pre- to post-testing on math concepts and problem-solving, math computational skills, math vocabulary (GMADE), and communication in math (BAM).

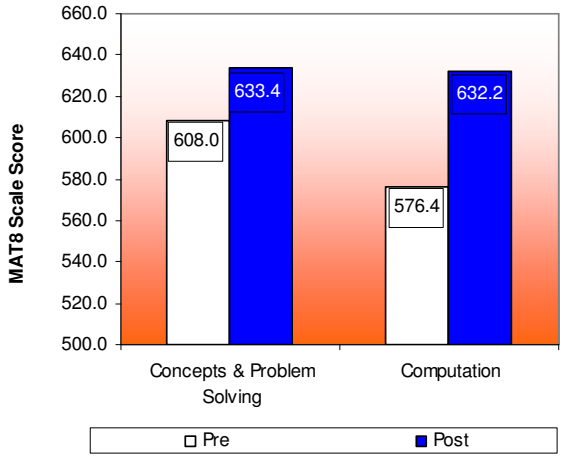
In order to facilitate interpretation of results, effect sizes for the gain scores were calculated and these estimates were translated into percentile gains[23]. For all assessments, results showed moderate to large effects (MAT8-Concepts and Problem-solving=.69; MAT8-Computation=.83; GMADE=.65; BAM=.86). Moreover, enVisionMATH students showed percentile gains of:

- 19 points on concepts and problem-solving,
- 33 points on computation,
- 15 points on math vocabulary, and
- 36 points on math communication.

[22] Detailed rationale on the statistical analyses performed is provided in the accompanying Technical Report, pages 5-8.

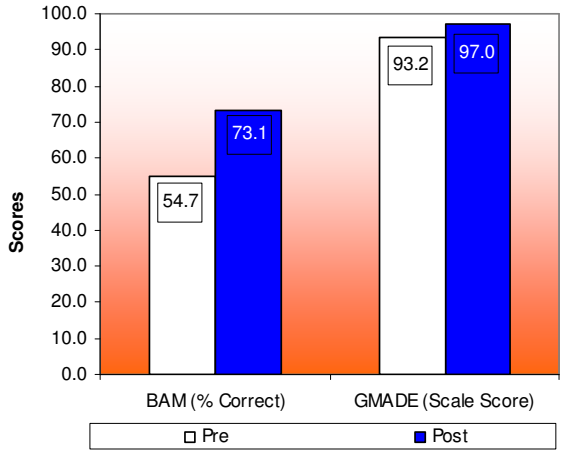
[23] The formula used for calculating gain score effect sizes was $d = M_1 - M_2 / \sigma_{pooled}$ (Rosnow and Rosenthal, 1996). Effect sizes can also be thought of as the average percentile standing of the average treated participant relative to the average untreated participant. As an example, an ES of 0.0 indicates that the mean of the treated group is at the 50th percentile of the untreated group. An ES of 0.8 indicates that the mean of the treated group is at the 79th percentile of the untreated group.

Figure 2. enVisionMATH Students' Math Performance at Pre and Post-testing: Metropolitan Achievement Test (MAT8)



- *There was significant improvement in enVisionMATH students understanding of math concepts and problem-solving and math computational skills. In these areas, students showed percentile gains of 19 and 33 points respectively.*

Figure 3. enVisionMATH Students' Math Performance at Pre and Post-testing: GMADE and Balanced Assessment in Math (BAM)



- *enVisionMATH students also showed significant gains in math communication and problem-solving (BAM) and in math vocabulary (GMADE). In these areas, enVisionMATH students showed percentile gains of 36 and 15 points respectively.*

enVisionMATH students' showed significant improvement in math concepts and problem-solving, math computation, math vocabulary, and communication in math over the course of the study.

Learning gains experienced by enVisionMATH students can also be seen in growth of percentile ranks[24] from the norm-based assessments. It is a general rule of thumb that if a student makes a year's growth for a year of instruction, then the percentile rank will remain the same. As shown in Table 8, the percentile rank grew more than would be expected in a typical academic year for math computation and math vocabulary. On the MAT8 Concepts and Problem-solving subtest, enVisionMATH students showed a typical amount of growth (i.e., equivalent to one year of instruction).

Table 8. enVisionMATH Students' Percentile Rankings at Pre and Post-testing

	Pre	Post
MAT8 Concepts and Problem-Solving	61st	61st
MAT8: Computation	54th	58th
GMADE: Math Vocabulary	36th	45th

- *enVisionMATH students had higher percentile rankings at post-testing than pre-testing on the MAT8 Computational subtest and the GMADE. Note that it is a general rule of thumb that if a student makes a year's growth for a year of instruction, then the percentile rank will remain the same (see results for MAT8 Concepts and Problem-solving). Therefore, enVisionMATH students grew more than would be expected over the course of a*

[24] Percentile ranks indicate the percent of students in the same grade in the norm (reference) groups who took the test at a comparable time and whose scores fall below a student's score. Since percentile ranks do not represent equal units, and since their interpretation is limited to the reference group from which they were derived, they are best used for reporting scores when position in relation to the reference group is of primary interest.

typical academic year as measured by the math computation and vocabulary tests.

Improvement on problem-solving and math communication is also evident on the BAM performance levels. Specifically, results showed that the enVisionMATH students moved to higher stages of mathematics understanding (see percents in green areas in Table 9). For example, 56.5% of students classified as "on target" at pretesting demonstrated a "strong" level of understanding at post-testing. In contrast, movement to lower stages was much less prevalent (see percents in red areas).

As well as enVisionMATH students showing significant gains on all three national assessments, enVisionMATH students showed more growth than would be expected over the course of a typical academic year as measured by the math computation and vocabulary tests.

Table 9. enVisionMATH Students' Level of Math Understanding (on BAM) at Pre and Post-testing

		Post			
		Limited Understanding	Beginning Understanding	On target	Strong (advanced) Understanding
Pre	Limited	15.0%	57.5%	18.8%	8.8%
	Beginning	1.3%	29.5%	38.7%	30.5%
	On target	0.0%	7.3%	36.3%	56.5%
	Strong (advanced)	0.0%	0.0%	5.0%	95.0%

- *Improvement among enVisionMATH students was evident on the Balanced Assessment of Mathematics Test. In particular, movement to higher levels of problem-solving and math communication skills was more prevalent than movement to lower stages. For example, 56.5% of students classified as "on target" during pretesting were at the strong or advanced level at post-testing.*

Do changes in math performance among enVisionMATH students vary by different types of students and levels of implementation?

Preliminary analyses were also conducted to examine if the enVisionMATH program was associated with improvements among students of various subgroups. Since the focus was to examine growth among different types of treatment student, these analyses only included the performance of enVisionMATH students in the following student subpopulations: females and males, special education and non-special education students, 2nd and 4th graders, minorities and non-minorities, students receiving free/reduced lunch and not, and students of various math levels. It should be noted that the sample sizes in the subgroups may be small and there are unequal sample sizes between those in the subpopulations and those not[25]. Therefore, with the caveat that these analyses are limited, this provides readers with preliminary, descriptive information on whether the program is associated with improvements among various subgroups. Figures 4 through 9 display the results for the various subgroups.

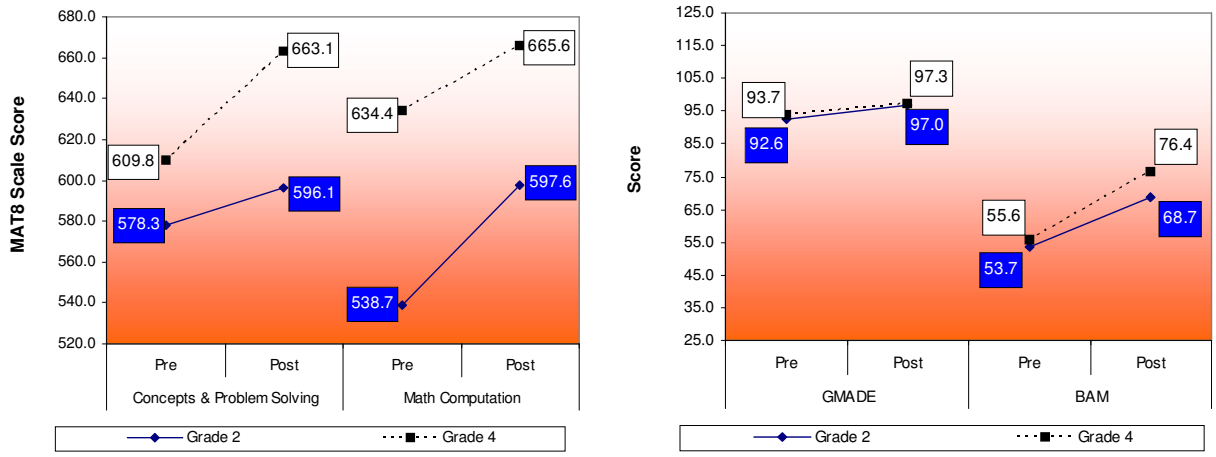
Results showed that enVisionMATH students in all subgroups significantly improved from pre- to post-testing on all math measures. That is, females and males, special education and non-special education students, 2nd and 4th graders, minorities and non-minorities, students receiving free/reduced lunch and those not, and students of various math levels all showed significant learning gains, $p<.05$.

In addition, differential growth rates were observed for all subgroups, with the exception of special education status. Specifically, 4th grade enVisionMATH students showed greater gains than 2nd grade students on the BAM. Males showed greater growth on the MAT8 Concepts and Problem-solving subtest than females, whereas females showed greater growth than males on the BAM. In addition, minority treatment students also showed more improvement on the MAT8 Computation subtest as compared to White students. Treatment students *not* receiving free-reduced lunch showed more improvement on the MAT8 Computation and BAM tests than treatment students receiving free-reduced lunch.

Overall, the enVisionMATH program worked just as well with 2nd and 4th graders, females and males, White and non-White students, special education and non-special education students, and students receiving free/reduced lunch and those not receiving this aid. That is, although a greater rate of improvement was demonstrated for certain subgroups of students, all enVisionMATH students showed significant gains in math concepts and problem-solving, math computation, math vocabulary, and communication in math.

[25] The reader is referred to Appendix B in the accompanying Technical Report for detailed statistics.

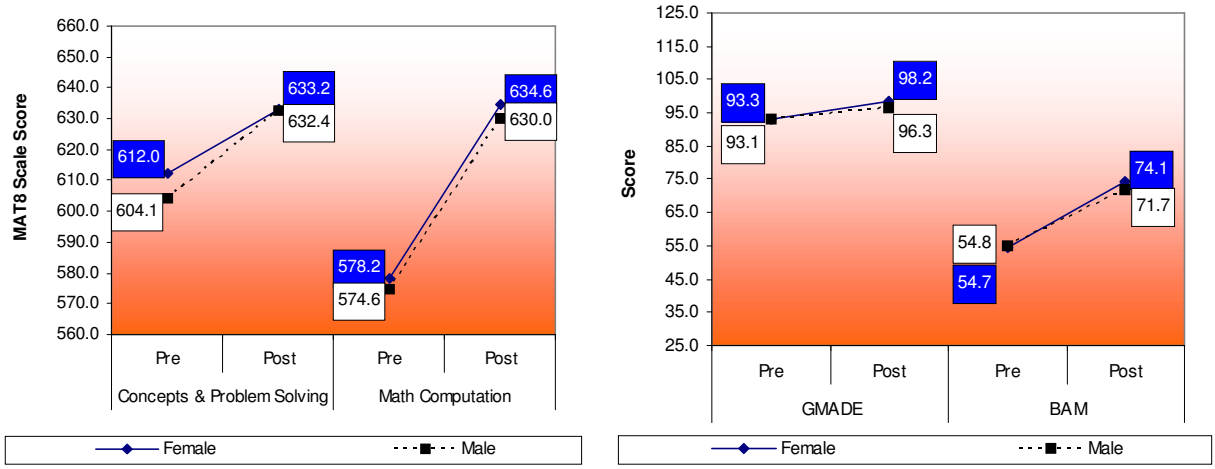
Figure 4. enVisionMATH Students' Assessment Performance Gains by Grade*



*NOTE: Since the MAT8 scale scores are developmental (i.e., increase with age/grade), analyses examining interaction of grade and gains on MAT8 were excluded.

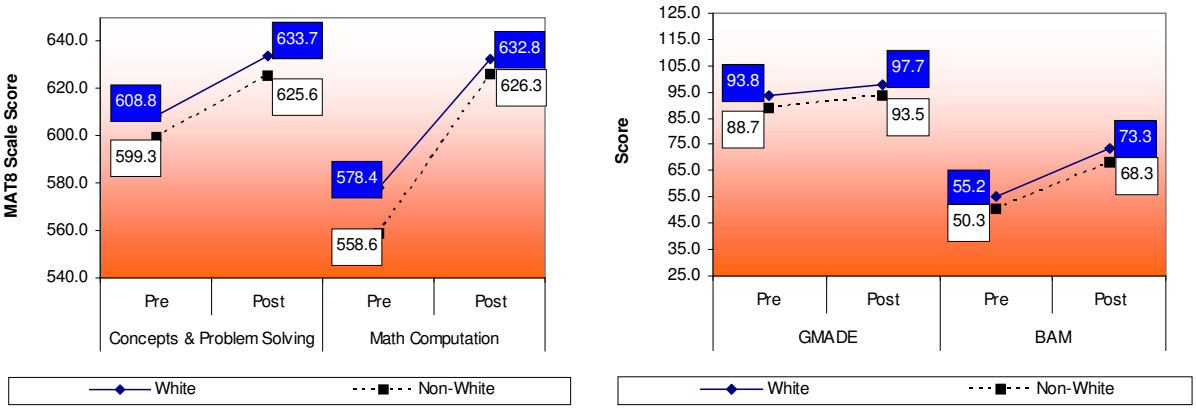
- *2nd and 4th grade enVisionMATH students showed significant improvement in math performance as measured by all assessments. In addition, 4th grade students showed more accelerated gains in problem-solving and math communication (as measured by the BAM) as compared to 2nd grade students.*

Figure 5. enVisionMATH Students' Assessment Performance Gains by Gender



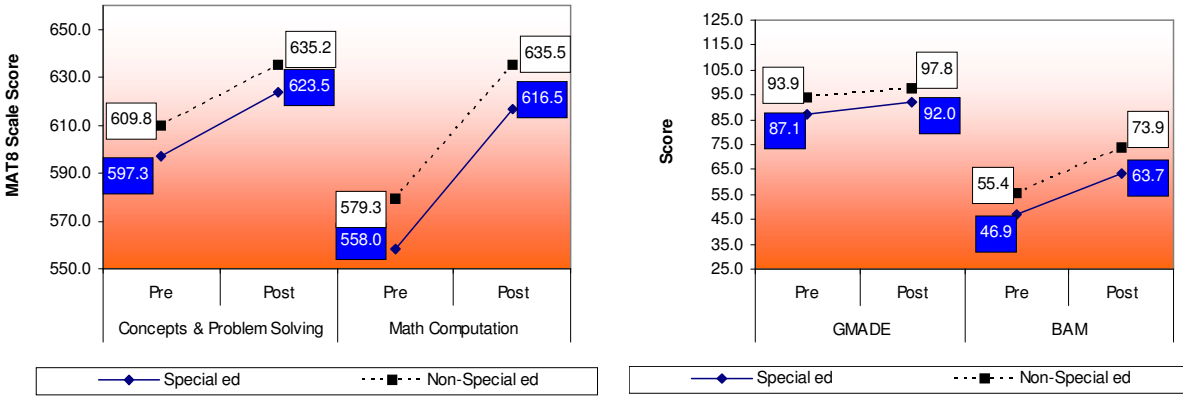
- *While significant improvement was observed among both males and females, males tended to show greater gains in Concepts and Problem-solving whereas females showed greater improvement on the BAM (measuring problem-solving and communication in math).*

Figure 6. enVisionMATH Students' Assessment Performance Gains by Ethnicity



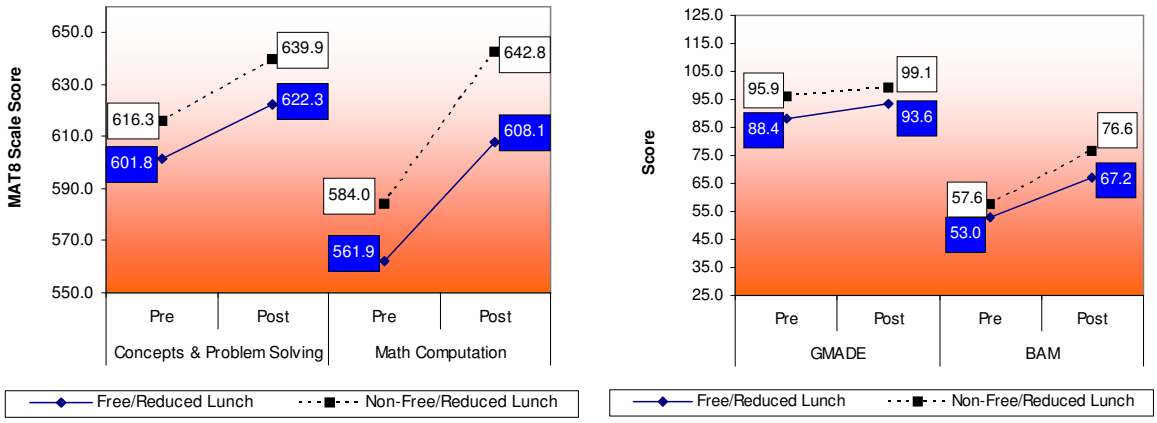
- *White and non-white enVisionMATH students showed notable gains on all assessments. Furthermore, non-white students showed higher rates of improvement on the math computation subtest as compared to White students.*

Figure 7. enVisionMATH Students' Assessment Performance Gains by Special Education Status



- *Special education and students not in special education showed significant gains in math performance across all measures after using enVisionMATH for one year. In addition, rates of improvement were similar across both types of enVisionMATH students.*

Figure 8. enVisionMATH Students' Assessment Performance Gains by Free/Reduced Lunch Status

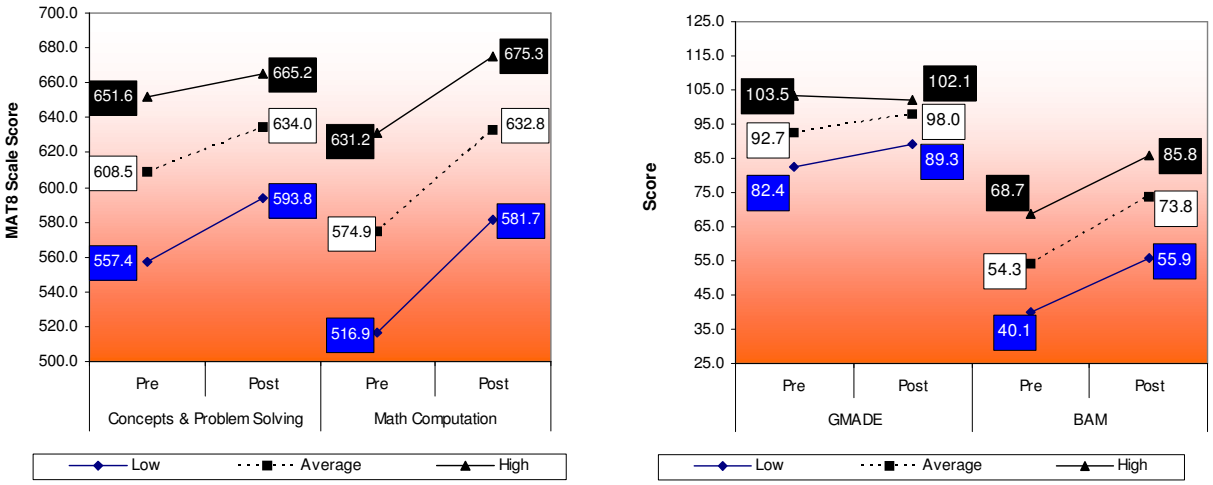


- *Significant improvement was observed among enVisionMATH students receiving free/reduced lunch and those not receiving this aid. In addition, results also showed more improvement in the performance of students not receiving free/reduced lunch on the Math Computation and BAM tests as compared to students receiving free/reduced lunch.*

In order to categorize students on initial math performance, the percentile rankings from the norm-referenced tests (MAT8 and GMADE) at pretest were used. Students who were at or below the 33rd percentile were classified at a low math level, students who were at or above the 66th percentile were classified as high, and the remaining students were classified as average. Comparisons were made between the three identified math levels. Results showed significant gains on all assessment measures, with the exception of high-level students on the GMADE. In addition, results showed accelerated gains by certain groups of students. Specifically, the change in performance from pre to post was greatest among low-level students, followed by average level students and then high level students as measured by the MAT8 subtests and GMADE. Note that this may be due to the fact that there is more room for improvement among low-level students as compared to higher level students. In contrast, on the BAM, which required students to respond to constructed-response items and with a focus on math problem-solving and communication, results showed that average students showed the greatest amount of change, followed by high level students and then low level students.

In general, results showed that enVisionMATH students at all levels of mathematics ability (low, average, and high) demonstrated significant gains in math performance. That is, enVisionMATH helps improve upon the existing ability levels of all types of learners so that they can reach a higher level of achievement.

Figure 9. enVisionMATH Students' Assessment Performance Gains by Math Ability Level



- *While students at low, average and high math ability levels showed improvement in performance, lower-performing students tended to show the greatest amount of growth, followed by average and then high-performing students, as measured by the MAT8 subtests and GMADE. It is important to note, however, that the observed relationship may be due to the fact that there is more room for improvement among the lower-level students as compared to higher performing students. In contrast, on the BAM, average students showed the most improvement, followed by high level students and then low level students.*

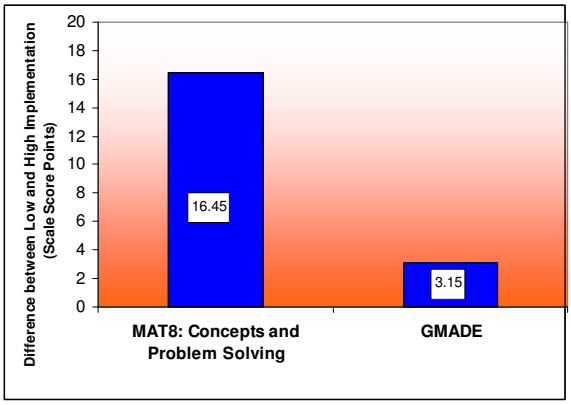
In addition to these analyses among subgroups of enVisionMATH students, exploratory multilevel analyses on the relationship between overall levels of enVisionMATH implementation of key program components and student posttest performance were conducted. These analyses provide preliminary information on whether low, moderate, and high implementation fidelity of enVisionMATH[26] components was associated with student performance.

Results showed that there was a significant relationship between overall enVisionMATH implementation levels and improved performance on the outcome measures[27]. Specifically, students whose teachers used the enVisionMATH program with high fidelity showed the greatest gains in math performance as measured by the MAT8 Concepts and Problem-solving subtest, t-ratio=2.93, ***p*=0.004**, and GMADE, t-ratio=2.33, ***p*=0.02**, as compared to students of teachers who did not use the program with high fidelity.

[26] See section on Fidelity of Implementation for how this categorization was determined.

[27] Detailed statistics are presented in the accompanying Technical Report.

Figure 10. Difference in Performance among Low and High Implementation enVisionMATH Classes



- *Students of teachers who implemented enVisionMATH with high fidelity showed an average of 16 more points on the MAT8 Concepts and Problem-solving subtest (2.7% gain) and 3 more points on the GMADE (3.3% gain).*

Preliminary analyses showed that students whose teachers implemented the major components of enVisionMATH with high fidelity showed greater improvement than students of teachers who implemented enVisionMATH with low fidelity and did not use the major program components on a regular basis.

These results suggest that teachers should aim to use the program fully as described in the implementation guidelines. This is because increased use of the program is associated with higher test scores.

The aforementioned analyses focused on the extent to which enVisionMATH is positively associated with student math performance (and it is). However, these analyses do not examine how students improved over time as compared to students using other math programs. The following section presents analyses of how the math performance of students exposed to enVisionMATH compares to the performance of students using other math programs.

How does mathematics performance differ between students who use enVisionMATH as compared to students who do not use this program?

Prior to discussing the results found, it is important to reiterate that there were a number of similarities between control and enVisionMATH classrooms. While differences did exist in the content taught, the majority of math concepts covered in the 2nd and 4th grade classes were consistent across control and enVisionMATH classes. This is to be expected since both treatment and control teachers within sites were following similar state standards and district curriculum guidelines. In addition, the instructional strategies and practices employed by treatment and control teachers were generally similar. Given this information and the fact that the duration of the study and exposure to the program occurred during *one* school year, small effect sizes are to be expected. Usually even with training provided, there is a learning curve for teachers in their first year of implementing a new program. Indeed, while 32.1% of treatment teachers indicated that it took about one month before they felt comfortable teaching enVisionMATH, 28.6% indicated they were not completely comfortable teaching the program until after one quarter of the school year and another 35.7% said it took half of the year before they were comfortable with the enVisionMATH program. Control teachers, however, had been using their programs for an average of 3-5 years. It can be expected that treatment teachers' comfort level and ability to use the program effectively will increase with more time and, as such, that

effect sizes observed will be larger[28]. Indeed, because the enVisionMATH study is a longitudinal study designed to take place over two school years, this will allow for *cumulative* student exposure to be examined over time in order to determine the long-term effects of the program, as well as sustainability of any observed effects.

RESULTS

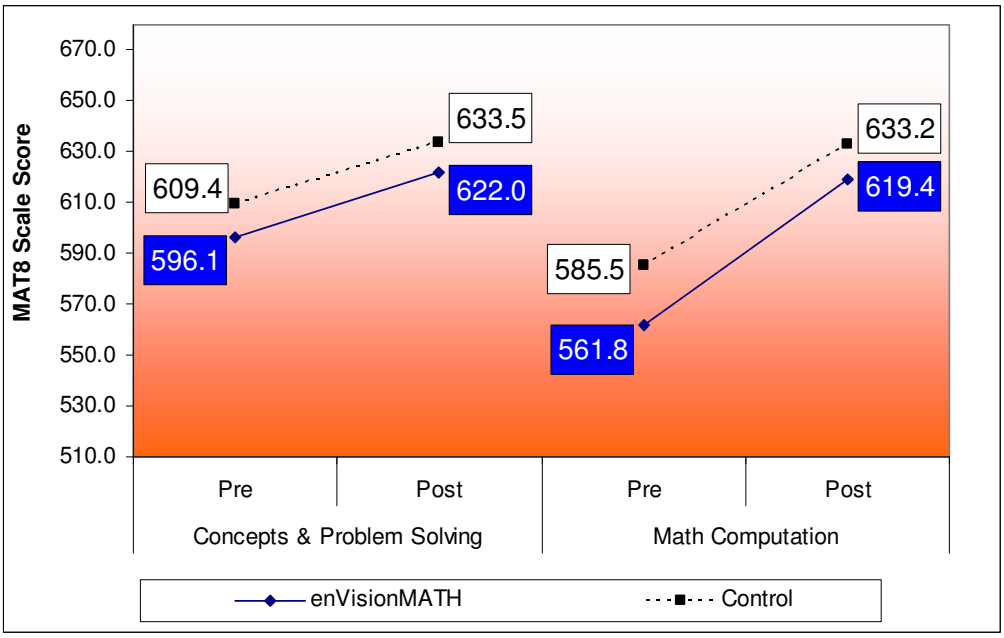
Multilevel models[29] were run to examine whether there was a significant difference in growth in performance between treatment and control students, as well as to account for statistical issues that can affect the validity of the results and to equate the groups on important variables (i.e., pretest, ethnicity, years of teacher experience, degree earned, discomfort in teaching mathematics, class climate and engagement, and school). Results showed significant differences between students using enVisionMATH and control students for the following measures: MAT8 Math Computation, t-ratio=9.79, ***p*=0.02**, *d*=0.21, GMADE, t-ratio=2.34, ***p*=0.01**, *d*=0.24 , and BAM, t-ratio=3.15, ***p*=0.01**, *d*=0.20. Specifically, students using enVisionMATH demonstrated greater improvement in math performance as compared to students not using the program in the areas of math computation, math vocabulary, and math problem-solving and communication. Note that significant effects were observed despite the fact that enVisionMATH students started out at a lower math level (see pretest scores); indeed, enVisionMATH students subsequently *surpassed* control students in the areas of math vocabulary (GMADE) and mathematics problem-solving and communication (BAM) during post-testing.

Results showed that students using enVisionMATH demonstrated significantly greater growth in math than control students across all three national math assessments used as outcome measures in this study (MAT8, GMADE, BAM).

[28] Note that treatment teachers participating in year 2 of this study were randomly assigned to conditions prior to the onset of this study. In particular, 3rd and 5th treatment teachers were provided with enVisionMATH materials and offered training during year 1 so that they could begin using the program prior to their participation in the study in year 2. Thus, the year 2 study teachers will already be familiar with program.

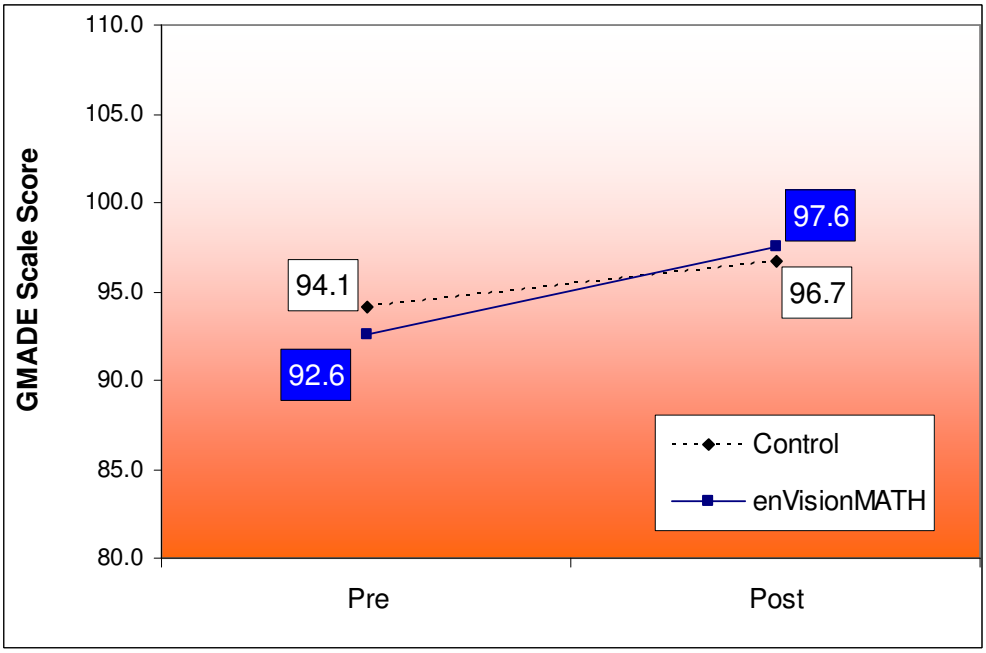
[29] Detailed information and statistics regarding these results are presented in the accompanying Technical Report.

Figure 11. Pre- and Post-test MAT8 Math Performance of enVisionMATH and Control Students



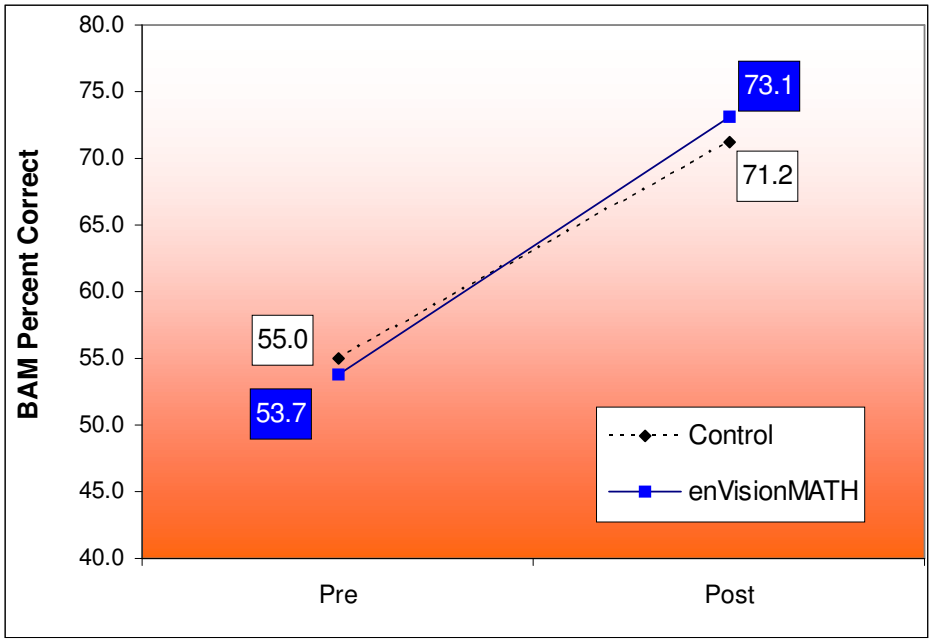
- *Growth analyses showed significantly greater gains on math computation (i.e., the blue line's slope is steeper). This means that enVisionMATH students showed greater rates of growth in math computation performance from pre- to post-testing as compared to control students.*

Figure 12. Pre- and Post-test GMADE Math Performance of enVisionMATH and Control Students



- *On the GMADE, a measure of math vocabulary understanding, enVisionMATH students showed significantly more improvement than control students. Note that while enVisionMATH students showed lower performance at pre-testing (though not significantly so), they subsequently surpassed control students by the end of the school year.*

Figure 13. Pre- and Post-test BAM Performance of enVisionMATH and Control Students



"A lot of information is covered in each lesson...and it makes the students think and apply the information in order to problem solve."- - 4th grade enVisionMATH Teacher

- *enVisionMATH students showed significantly greater gains in math problem-solving and communication as compared to control students. Although enVisionMATH students showed lower performance at pre-testing (though not significantly so), they subsequently surpassed control students and showed higher test scores at post-testing.*

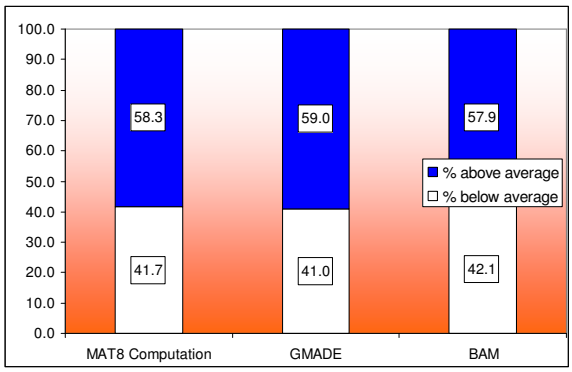
Elementary students who used enVisionMATH showed significantly greater gains in math computation, math vocabulary, and math problem-solving and communication as compared to students who used other math programs.

Effect size is a commonly used measure of the importance of the effect of an intervention (in this case, enVisionMATH). Given the similarity in treatment and control classrooms and the fact that this was the first year of the two-year study, small (.20) effect sizes were expected. Indeed, the effect sizes obtained can be classified as small (d=.21 for Math Computation, d=.24 for GMADE, and d=.20 for BAM). While small, these effect sizes are all statistically significant and meaningful – especially given the short duration of the study and the applied settings in which the study was undertaken. Given that this study will continue into a second year (2008-2009 school year), larger effects could be expected if such trends continue and students are exposed to enVisionMATH for even longer periods of time.

In order to better understand the effects observed as a result of exposure to enVisionMATH, effect sizes can be translated to the percent of treatment students that can be expected to be *above* the average of the control group (see blue part of bar in Figure 14). As shown, 58%, 59%, and 58% of students using the enVisionMATH program are more likely to have scored above the average of control students in the areas of math vocabulary, problem-solving and communication of math, and math computation, respectively.

It should be noted that the WWC calculates an improvement index which represents the difference between the percentile rank of the average student in the intervention condition (i.e., enVisionMATH) 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. Using the aforementioned effect sizes, the improvement index for this study can be calculated to be approximately +8, a noteworthy figure.

Figure 14. Percent of enVisionMATH Students Above and Below Average Relative to Control Students



- *Results show that 58%, 59%, and 58% of enVisionMATH students scored above the average control student on math computation, math vocabulary (GMADE), and math problem-solving and communication (BAM), respectively.*

While effects were small, they are all significant and meaningful – especially given the short duration of the study and the applied settings in which the study was undertaken. Given that this study will continue into a second (2008-2009 school year), larger effects could be expected if such trends continue and students are exposed to enVisionMATH for even longer periods of time.

"I love it. I can see where it helps the children to have the higher level thinking skills, and the way the questions are asked and the math is presented is not typical…it is more like the way I feel they will see it on end of grade tests or standardized test. It goes hand in hand and uses the terms they will have on the test." - 2nd Grade enVisionMATH Teacher

Do effects on student math performance between enVisionMATH and control students differ across types of students or settings?

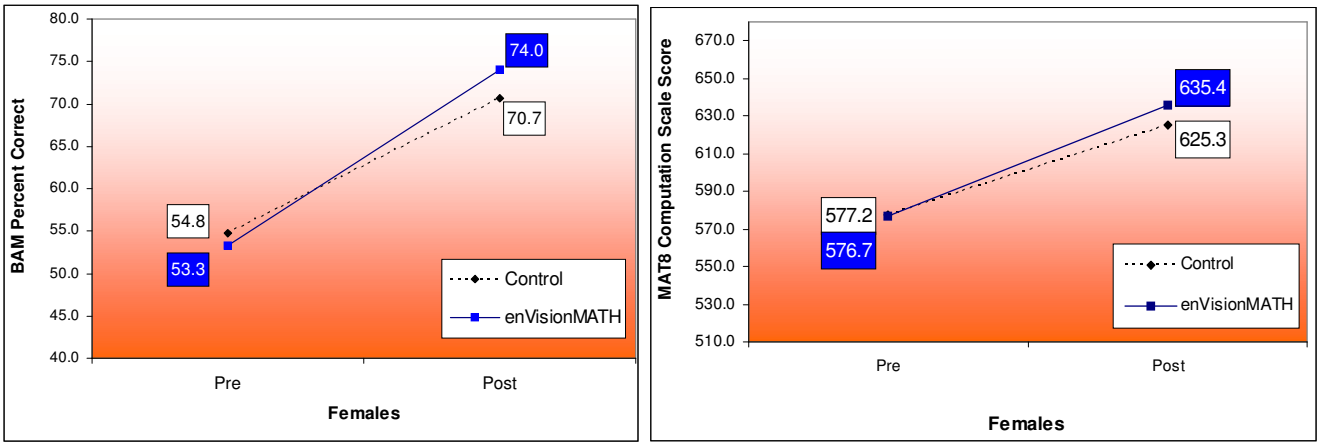
To examine if there were differences in performance between different subgroups of enVisionMATH and control students, subgroup effects were analyzed via multilevel modeling. Specifically, differences between enVisionMATH and control students in the following subgroups were examined: grade, gender, ethnicity, free/reduced lunch status, special education status, math ability level, and school. As previously noted, multilevel models account for statistical issues that can affect the validity of the results. Furthermore, it is important to view these analyses as exploratory[30]. Significant subgroup differences are discussed in the following sections.

RESULTS BY STUDENT SUBPOPULATIONS

Results showed a significant difference between enVisionMATH students and control students in the following subgroups: 4th grade, minorities, and females. These results are shown in Figures 15-17. Specifically, females and 4th grade enVisionMATH students showed greater math gains on the MAT8 Math Computation, t-ratio=10.16, ***p*=0.02** and t-ratio=17.66,***p*=0.03**, and BAM tests, t-ratio=3.25, ***p*=0.01** and t-ratio=3.22, ***p*=0.05**, as compared to control students. This suggests that the enVisionMATH program has a more positive impact on the computational, and math problem-solving and communication skills of females and 4th graders as compared to other math programs.

In addition, while minority students in the control group showed greater gains on the MAT8 Concepts and Problem-Solving subtests, t-ratio=-19.87, ***p*=0.01**, results also showed that enVisionMATH minority students showed more improvement on the GMADE, t-ratio=6.84, ***p*=0.01**. Due to the inconsistency in these results among minority students, results should be viewed with caution.

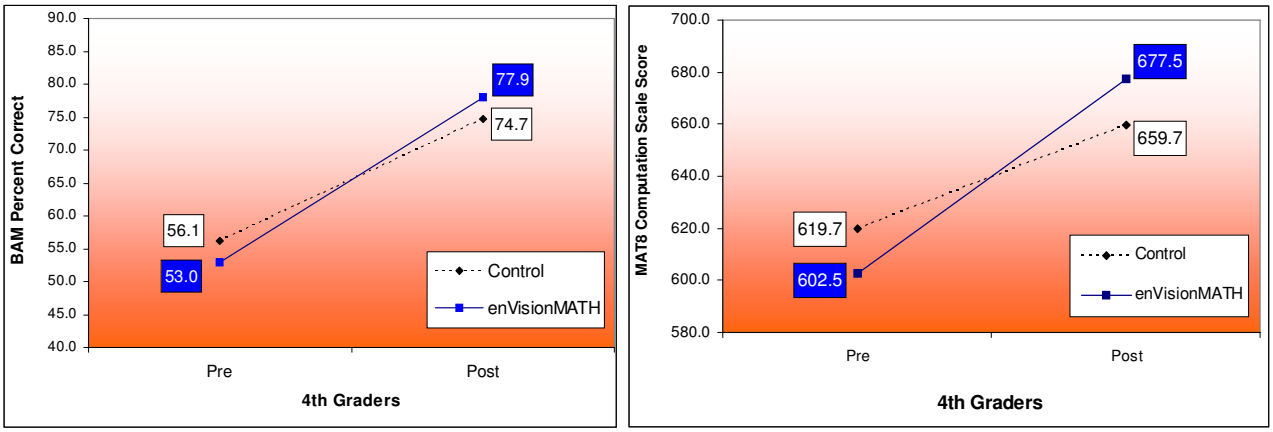
Figure 15. Pre-Post Math Performance of enVisionMATH and Control Students who are Female



- *Females who used enVisionMATH showed more significant growth in math problem-solving and communication, and in math computation as compared to females who used other math programs.*

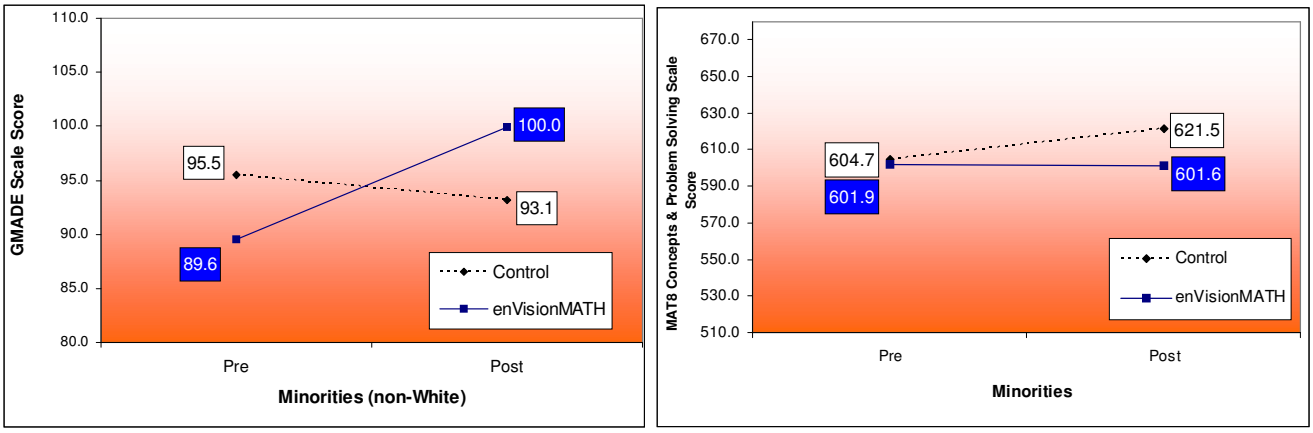
[30] Detailed information on why this is exploratory and non-causal and statistics regarding these results are presented in the accompanying Technical Report.

Figure 16. Pre-Post Math Performance of enVisionMATH and Control Students who are 4th Graders



- *Similarly, results showed that 4th grade enVisionMATH students displayed more accelerated gains in math problem-solving and communication, and in math computation. Indeed, while 4th grade enVisionMATH students started out at a lower math performance level than 4th grade control students, they later surpassed 4th grade control students at post-testing.*

Figure 17. Pre-Post Math Performance of enVisionMATH and Control Students who are Minorities



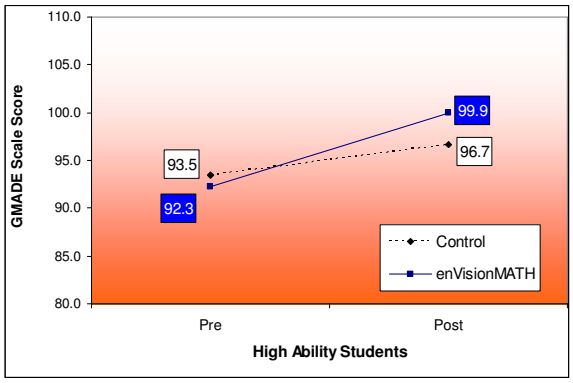
- *Results by minority students showed inconsistencies. While minority control students showed more significant gains in their performance on concepts and problem-solving, enVisionMATH minority students showed greater gains in the area of math vocabulary (GMADE). Given these inconsistencies in results, these findings should be interpreted with caution and no definitive conclusions should be made with regard to the impact of enVisionMATH on minorities.*

Examination of subgroup differences showed that enVisionMATH students who are females as well as those who are in 4th grade had greater math gains in the areas of math problem-solving and communication, and computation as compared to control students in these two subgroups. In contrast, results among minorities were inconsistent and therefore, findings should be interpreted with caution.

RESULTS BY MATH ABILITY

Significant differences were also observed among high math ability students. As previously noted, students were classified into math levels depending on their percentile rankings during pretesting. Students who scored at or above the 66th percentile were classified as high math ability students. Results showed that high level math students who used enVisionMATH demonstrated greater improvement on the GMADE as compared to high math level control students, t-ratio=3.27, ***p*=0.003** (see Figure 18).

Figure 18. Pre-Post Math Performance of enVisionMATH and Control Students who are of High Math Ability



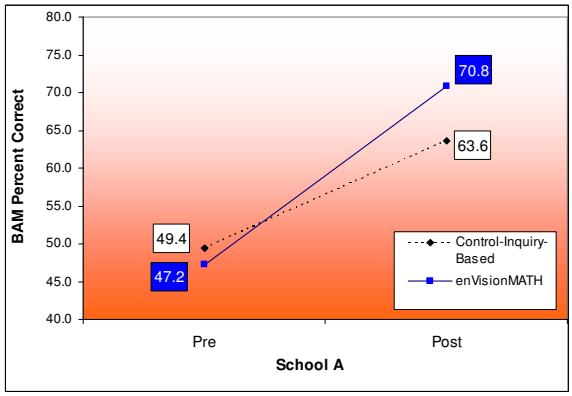
High performing enVisionMATH students showed more accelerated gains in math vocabulary as compared to high performing students using other math programs.

RESULTS BY TYPE OF CONTROL PROGRAM

Comparisons between enVisionMATH and Inquiry-Based Math Control Program

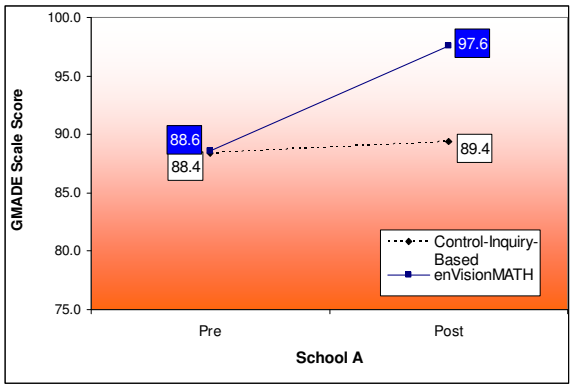
Analysis were performed to examine if there were differences between students who used enVisionMATH, which consists of a blend of traditional and inquiry-based approaches, and those who were exposed to a purely inquiry-based approach. Specifically, multilevel analysis were conducted to examine if differences existed between enVisionMATH and control students at school A. Results showed that enVisionMATH students had greater gains as compared to control students using a solely inquiry-based approach as measured by the BAM, t-ratio=7.95, ***p*=0.001**, and GMADE, t-ratio=9.47, ***p*=0.003** (see Figures 19 and 20).

Figure 19. Pre-Post Math Performance of enVisionMATH and Control Students Using an Inquiry-Based Approach: BAM



- *Control students at school A who were exposed to a purely inquiry-based math program showed significantly less growth from pre- to post-testing as compared to enVisionMATH students. Indeed, enVisionMATH students showed accelerated gains in math communication and problem-solving as compared to the control students.*

Figure 20. Pre-Post Math Performance of enVisionMATH and Control Students Using an Inquiry-Based Approach: GMADE



- *Similarly, while control students who used the inquiry-based math program showed very little gains on math vocabulary, enVisionMATH students demonstrated much more significant gains.*

Analysis conducted to examine if there were differences between enVisionMATH students and control students who were exposed to a purely inquiry-based approach (school A) showed that enVisionMATH students had significantly greater gains in math communication and problem-solving, and math vocabulary as compared to the control students.

Comparisons between enVisionMATH and Traditional Basal Math Control Programs

Numerous differences were also observed between enVisionMATH and control students at schools using traditional basal math programs. In particular, comparisons were made between enVisionMATH students and control students at schools B, C, D, E, F, and H, all of which used traditional basal math programs. Note that while all programs can be classified as traditional basal math programs, the control math programs were also distinct in a number of ways (see Appendix C for detailed information on the control programs).

Statistically significant differences are displayed in Figures 21-23. Specifically, results showed the following:

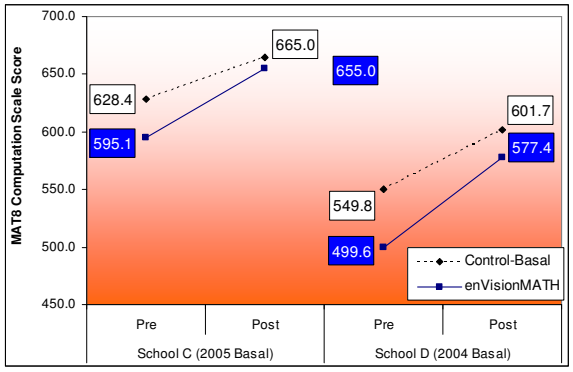
- enVisionMATH students attending school C showed more positive effects on the MAT8 Computation , t-ratio=3.69, ***p*<0.001**, and BAM tests, t-ratio=3.30, ***p*<0.01**, than control students; this school used a traditional basal math program (2005) that was very similar to enVisionMATH, incorporating many similar program elements. The enVisionMATH program, however, places more emphasis on integrating both traditional and investigative learning opportunities.
- Greater improvement was observed for enVisionMATH students at school D on the MAT8 Computation subtest, t-ratio=2.21, ***p*=0.03**; the control program was a 2004 traditional basal math program which was also similar to enVisionMATH in that the program is broken out into a larger number of shorter, more specific chapters. The scope and sequence, while is similar to enVisionMATH, was also more specific to state standards. Additionally, the program pedagogy emphasized conceptual understanding, problem-solving and reasoning skills more than basic fact practice and computation.
- Accelerated gains among enVisionMATH students were also observed at school F on the GMADE, t-ratio=2.36, ***p*=0.02**. The control program was a newer 2007 traditional basal math program that places a lot of emphasis on developing concepts and skills over time in a variety of contexts as well as teaching students that there are multiple methods and strategies for problem-solving, though there is a lack of manipulatives and hands-on components.
- More significant growth on the GMADE, t-ratio=2.24, ***p*=0.03**, was also observed among enVisionMATH students at school B as compared to control students. This school used an older math program (1998) that, like enVisionMATH, incorporated the use of visual models throughout the lesson to help

students grasp concepts and associate them to real life. The primary difference between the two programs is that the control textbook had a much more traditional chapter set-up, with fewer chapters covering more content.

- The only negative effect whereby control students outperformed enVisionMATH students on the BAM was observed at school E, t-ratio=-3.93, $p<0.001$. The control program is the same as that used at school D. Of note is that the 2nd grade treatment teacher showed a low level of implementation and only completed 65% of the enVisionMATH program. As previously noted, high implementation fidelity is associated with increased math performance. Therefore, this difference may be the result of the low level of implementation exhibited by the enVisionMATH teacher.

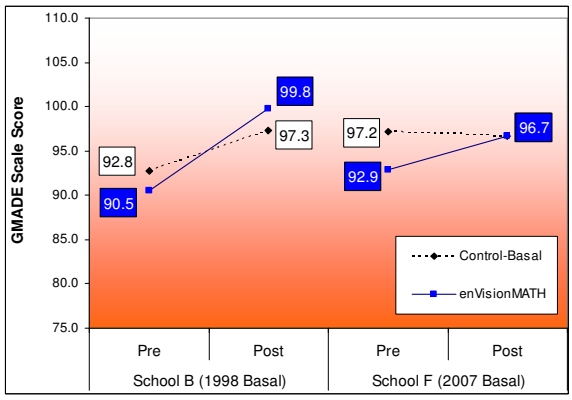
Given that schools, with the exception of schools D and E, used distinctive traditional basal math programs, the consistency of the majority of results across different schools, curricula, and measures, lends credence to the conclusion that enVisionMATH has a more positive impact on student math performance as compared to other traditional basal math programs.

Figure 21. Pre-Post Math Performance of enVisionMATH and Control Students Using Traditional Basal Math Programs: MAT8 Computation



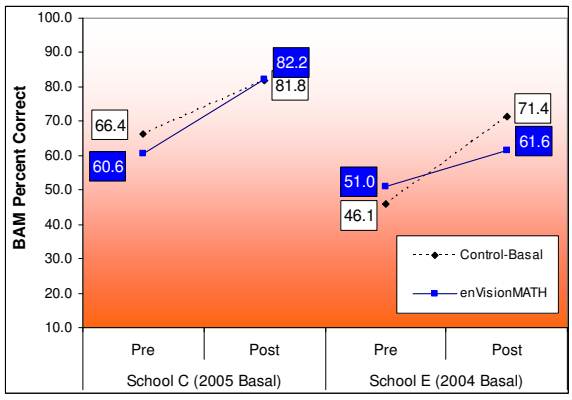
- *enVisionMATH students showed more significant gains in math computation than control students at schools C and D. Control students at these schools used distinct traditional basal math programs.*

Figure 22. Pre-Post Math Performance of enVisionMATH and Control Students Using Traditional Basal Math Programs: GMADE



- *Similarly, enVisionMATH students showed more accelerated growth on math vocabulary as compared to control students at schools B and F. Indeed, while enVisionMATH students started out at a lower level, they subsequently surpassed or caught up to the control students. These schools also used distinctive, yet traditional basal math programs; in addition, the program used at school B was the oldest used in the study while the control program used at school F was the newest (aside from enVisionMATH).*

Figure 23. Pre-Post Math Performance of enVisionMATH and Control Students Using Traditional Basal Math Programs: BAM



- *enVisionMATH students also showed more significant gains on math problem-solving and communication as compared to control students using a traditional basal math program at school C. In contrast, control students at school E showed more accelerated growth than enVisionMATH students. However, this may be due in part to the low*

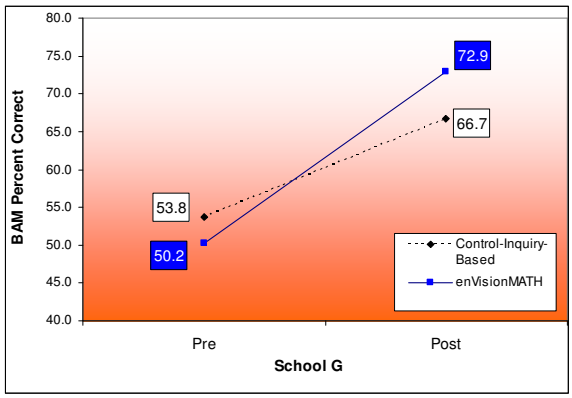
level of implementation by one of the two enVisionMATH teachers.

Overall, results show that enVisionMATH students outperformed control students who used a traditional basal math program in the areas of math problem-solving and communication, math vocabulary, and math computation.

Comparisons between enVisionMATH and Mixed, Teacher Created Control Program

Control teachers at school G used a mixture of math programs as well as teacher-created materials for their math instruction. Therefore, comparisons were made between students at school G who used enVisionMATH versus those in control classes whereby teachers used numerous resources, mixing traditional and inquiry-based activities. Results showed that enVisionMATH students had greater gains in math problem-solving and communication than control students, $_{t\text{-}}$ ratio=2.72, ***p*=0.007**, see Figure 24.

Figure 24. Pre-Post Math Performance of enVisionMATH and Control Students Using a Mixture of Math Programs: BAM



- *enVisionMATH students at school G showed more significant gains on math problem-solving and communication compared to control students whose teachers used various math programs and other teacher-created materials for math instruction.*

In summary, the positive effects obtained on the enVisionMATH program were observed across a number of different schools who used a variety of different control programs. enVisionMATH students performed significantly better than control students who used programs that were purely investigative and inquiry-based as well as students who used more traditional basal math programs. This consistency in findings across different curricula, schools, and measures lends credence to the conclusion that enVisionMATH positively impacts student math knowledge and skills.

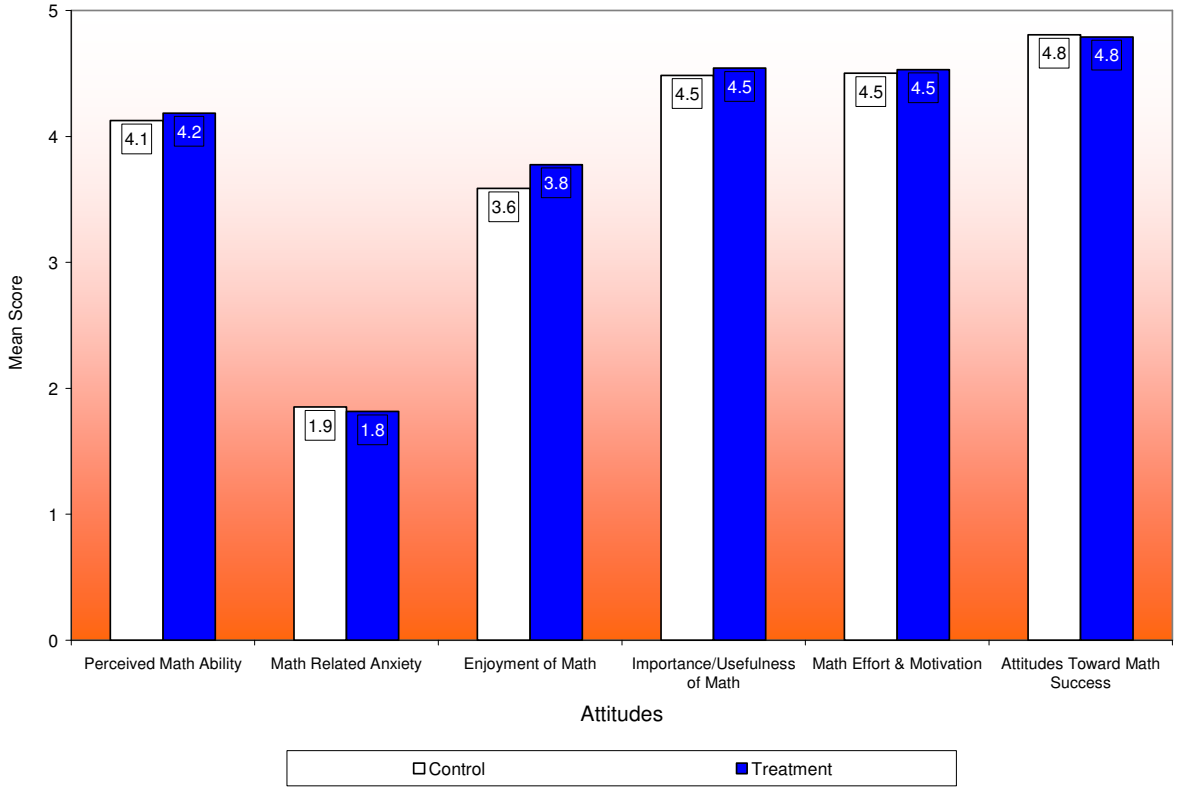
Does participation in enVisionMATH result in other positive outcomes?

While the primary focus of the enVisionMATH program is to improve students' math understanding and skills, the program incorporates a number of program components that may have an effect on other important aspects of math education, including affective attitudes. Measures were included in the RCT to explore whether use of the enVisionMATH was associated with changes in student attitudes towards math as well as changes in teacher practices and attitudes.

STUDENT ATTITUDES

Comparison of data collected on math-related student attitudes showed a significant effect for enjoyment of math, $F(1, 1036) = 8.58$, $p = .003$, $d=.21$. Notably, enVisionMATH students indicated greater enjoyment for math in the Spring survey, after controlling for pre-survey attitudes. No other significant differences were observed between enVisionMATH and control students on perceived math ability, math-related anxiety, importance/usefulness of math, math effort and motivation, and attitudes toward success in math, $p>.05$, see Figure 25.

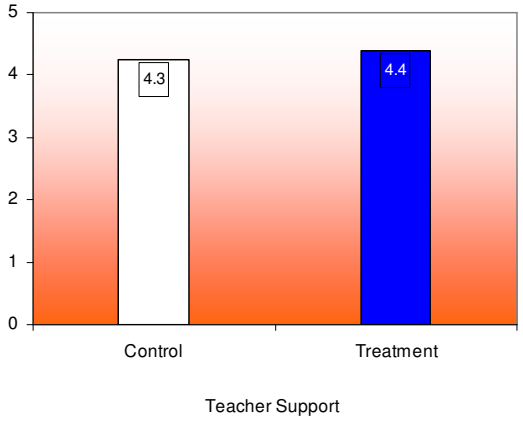
Figure 25. Student Math-Related Attitudes* by Group



*With exception of math-related anxiety, higher scores indicate more positive attitudes. Based on scale of 1-5.

In addition, a significant difference was found between enVisionMATH and control students on the measure for teacher support, $F(1, 1036) = 9.553$, p = .002, d=.16. Specifically, enVisionMATH students indicated experiencing greater support from their teachers as compared to control students, after controlling for pre-attitudes (see Figure 26). No differences were observed for perceived parental support, or school-related attitudes, $p>.05$.

Figure 26. Student Perceptions of Teacher Support



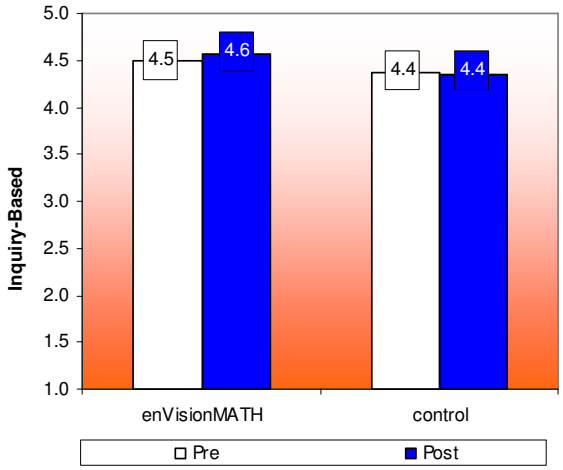
*Higher scores indicate more positive attitudes. Based on scale of 1-5.

While enVisionMATH and control students had similar positive attitudes on various areas (e.g., importance of math), significant differences were also observed. Results showed that enVisionMATH students enjoyed math more and perceived greater teacher support as compared to control students.

TEACHER ATTITUDES

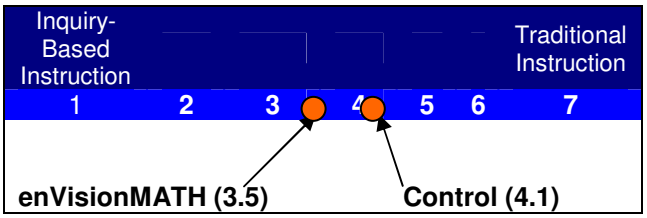
Teachers were asked about their attitudes regarding student learning, particularly their beliefs on inquiry-based instruction versus traditional basal methods of instruction. Comparison of their perceptions of both methods of instruction from the Spring survey revealed a marginally significant difference such that enVisionMATH teachers had a greater leaning towards inquiry-based approaches as compared to control teachers $t(51)=1.72$, $p=.091$, d=.45. It should be noted that no differences were observed between the two groups during the pre-survey on this measure.

Figure 27. Teacher Attitudes about Inquiry-based Instruction



These results are supported by findings from the Spring survey in which teachers were asked to classify their teaching style on scale of 1-7, with 1 being inquiry-based and 7 being traditional. As shown in Table 9 below, enVisionMATH teachers leaned more toward an inquiry-based approach as compared to control teachers, t(50)=2.122, ***p*=.04**.

Table 9. Self-Reported Inquiry versus Traditional Instructional Approach

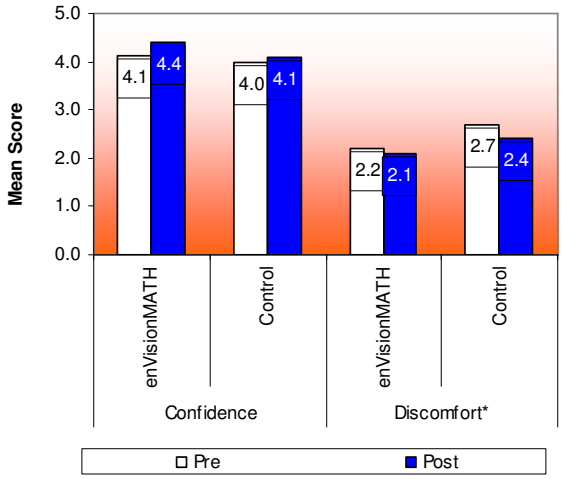


Teachers were also asked how well they felt prepared for various types of math activities and implementation of math strategies. While no differences were found in responses patterns between enVisionMATH and control teachers on the pre-survey, post-survey analyses revealed a significant positive effect. Overall, enVisionMATH teachers felt more prepared than their control counterparts, t(51)=3.49, ***p*=.001** d=.94. Specifically, enVisionMATH teachers indicated they felt more prepared to do the following:

- Provide hands-on, concrete experience before introducing abstract concepts, t(50)=1.83, ***p*=.07**
- Help students develop communication skills related to math (e.g., have students explain how they arrived at math solutions), t(51)=2.47, ***p*=.02**
- Engage students in applications of mathematics in a variety of contexts, t(51)=3.05, ***p*=.004**
- Teach using hands-on investigative activities related to math, t(51)=2.26, ***p*=.03**
- Use informal questioning to assess student understanding, t(51)=1.88, ***p*=.07**
- Help students with problem-solving skills, t(51)=2.25, ***p*=.03**
- Help students engage in mental math (e.g. estimation, thinking about the solution, etc.), t(51)=2.58, ***p*=.01**
- Teach different methods to solve mathematics problems, t(51)=2.08, ***p*=.04**
- Employ a variety of assessment techniques to gauge my students' level of understanding, t(51)=2.67, ***p*=.01**
- Utilize various classroom management techniques to maintain a productive learning environment, t(51)=2.01, ***p*=.05**

Teachers were also asked about their level of confidence and discomfort in teaching mathematics. Again, results showed a significant difference such that enVisionMATH teachers reported a greater level of confidence to teach math as compared to control teachers, p = .023, d=.63. No significant differences were observed on the discomfort measure[31].

Figure 28. Teacher Confidence and Discomfort to Teaching Math



*For confidence, higher scores indicate more confidence. For discomfort, higher scores indicate more discomfort. Based on scale of 1-5.

[31] Note that control teachers reported higher levels of discomfort in teaching mathematics compared to treatment teachers on the pre-survey, p = .042. To account for these differences on the pre-survey, post-survey analyses were conducted controlling for pre-survey levels of discomfort.

Results pertaining to their level of NCTM knowledge and perceptions about the usefulness of math revealed no differences between enVisionMATH and control teachers. However, results did show a significant increase in enVisionMATH teacher's knowledge of NCTM standards and focal points from fall to spring, $p<.05$. No such differences were observed among control teachers.

enVisionMATH teachers noted that they were more prepared to use various math practices and strategies as compared to control teachers. In addition, they tended to actually engage in more varied math activities. enVisionMATH teachers also reported greater confidence to teach math as compared to control teachers and showed a significant increase in their perceived knowledge of NCTM standards and focal points.

CLASSROOM PRACTICES

Teachers were also asked about the typical classroom practices they incorporated into their instructional day. While comparisons between enVisionMATH and control teachers on pre-survey measures revealed no significant differences, analyses of the Spring survey showed significant differences. Notably, enVisionMATH teachers tended to have their students engage in a greater variety of activities, $t(51)=2.95$, $p=.01$, $d=.80$. Follow-up analyses were conducted on the individual items comprising the Student Activities measure to better understand how enVisionMATH and control teachers differed. Specifically, enVisionMATH teachers noted they were more likely to have students do the following:

- Answer textbook/worksheet questions in-class, $t(51)=1.98$., $p=.05$
- Read a story that teaches a math concept or includes math-related ideas, $t(51)=1.76$, $p=.08$
- Use mathematical concepts to solve real-world problems, $t(51)=1.76$, $p=.09$
- Engage in problem-solving exercises or activities, $t(51)=2.07$, $p=.04$
- Engage in mental math (e.g. estimation, thinking about the solution, etc.), $t(51)=2.21$, $p=.03$

No differences in other teacher practices were observed.

In summary, results showed that enVisionMATH had positive effects on student and teacher attitudes, and teacher practices. In particular, enVisionMATH students displayed greater enjoyment of math and perceived more teacher support as compared to control students. In addition, enVisionMATH teachers experienced greater confidence in teaching math, were more prepared to carry out various mathematics activities, and in fact, tended to engage in a greater variety of mathematics activities and strategies as compared to control teachers. These findings are noteworthy because even with the small sample size (n=52) and associated low power, these findings were significant and effect sizes were moderate to large (d=.45 to .94). In addition, following one year of enVisionMATH use, teachers tended to lean more towards an inquiry-based approach. That said, their self-rating was still somewhat in the middle range (not completely inquiry-based nor traditional); this is to be expected among enVisionMATH teachers as it attempts to blend both types of pedagogies.

What did users of enVisionMATH program think about the program?

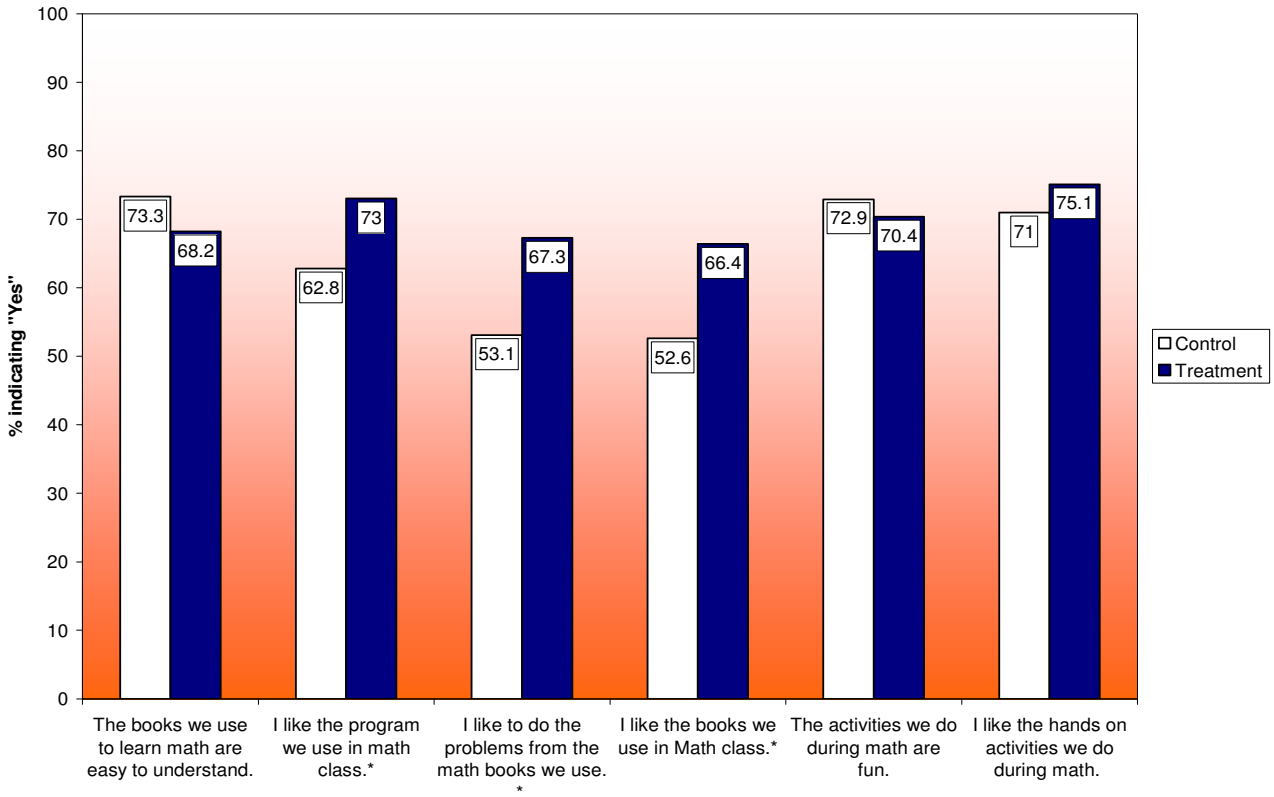
"{Pearson} has done really well with this series. I am very pleased with enVisionMATH. I have enjoyed teaching math using these materials very much." -- 2nd Grade enVisionMATH teacher

"I really like the hands-on activities, it provides a good introduction. The Visual Learning Bridge always reinforces what they are learning. {The program} also has a really nice layout." -- 2nd Grade enVisionMATH teacher

Teachers assigned to use the enVisionMATH program during the first year of the study overwhelmingly liked the program. Information gathered throughout the first year of the study showed that teachers felt the program helped them to teach more effectively.

Students in enVisionMATH classrooms also enjoyed using the program. Analysis of student surveys showed that students using enVisionMATH had more positive perceptions about the materials they used for math than control students. Significant differences were observed such that enVisionMATH students rated their program higher than control students (see Figure 29). As shown, 73% of enVisionMATH students noted that they liked the program used in math class as compared to 62.8% of control students.

Figure 29. Student Attitudes about their Assigned Math Program



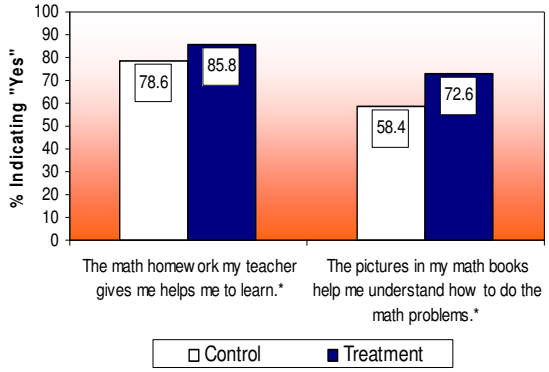
*Significantly different at the p<.05 level.

This is also confirmed by the item "I like the books we use in Math class", whereby 66.4% of enVisionMATH students agreed as compared to 52.6% of control students. Additionally, a higher percentage of enVisionMATH students (67.3%) noted they liked the problems from their textbooks as compared to control students (52.6%).

"The program is very engaging and the kids love it, especially the Interactive Learning Activity." --2nd grade enVisionMATH teacher

Positive results were also observed with regard to student perceptions of the extent to which the enVisionMATH program helped them to learn math. As shown in Figure 30, 85.8% of enVisionMATH students agreed that the homework assigned in their math class helped them to learn. Moreover, 72.6% agreed that the pictures in their math book helped them to understand how to do math problems compared to 58.4% of control students.

Figure 30. Student Attitudes about the Helpfulness of their Assigned Math Program

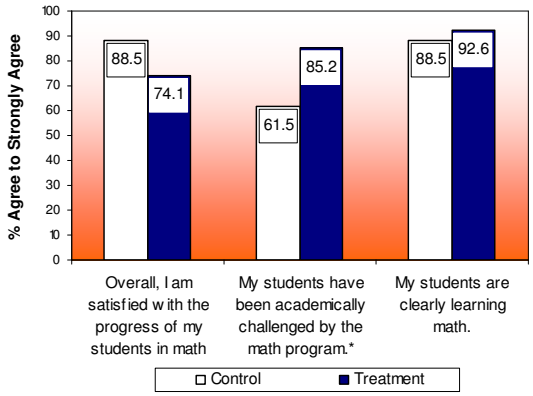


*Significantly different at the $p<.05$ level.

"I have one student that I worked with in third grade that would be classified as an 'at risk student' and that particular student has thrived with enVisionMATH; something has clicked and just really worked well." -- 4th Grade teacher

Teachers also felt that the enVisionMATH program was effective in teaching their students math. For instance, 92.6% of enVisionMATH teachers agreed that their students were clearly learning math (see Figure 31), and 74.1% indicated that they were satisfied with the progress their students were making. Figure 31 also shows that 85.2% of teachers agreed that their students were academically challenged by the program, compared to only 61.5% of control teachers.

Figure 31. Teacher Attitudes about their Math Program's Effect on Student Learning



*Significantly different at the $p<.05$ level.

"A lot of students who did not have good math skills in the beginning have seen some nice gains." -- 2nd Grade enVisionMATH Teacher

{The program} is more visual to them. {Students} can pull the numbers apart. During the test I could see them using their fingers for the tens spot. I can see them breaking those numbers apart. They were pretending they were busting the numbers apart. They have a better number sense. They can actually see them." -- 4th Grade enVisionMATH Teacher

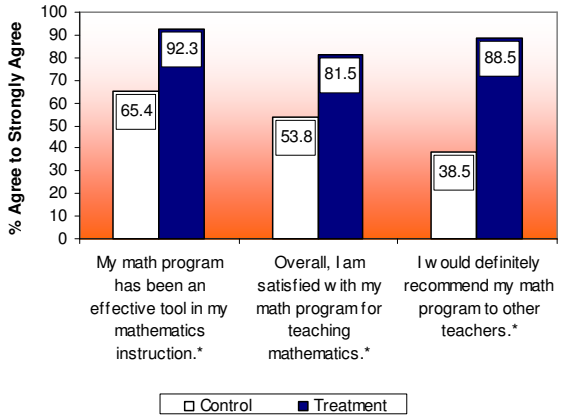
Teachers overwhelming indicated that they liked the enVisionMATH program better than the program they had used the previous year. All but one participating

teacher indicated that they liked the program "better" or "much better" than their past program. enVisionMATH teachers noted that they felt they had more flexibility in teaching with this program and appreciated the abundant materials and resources available with the enVisionMATH program. One teacher noted that she liked this program much more than her previous program because "it is broken into smaller pieces and makes the topics easier for the students to understand."

"Overall I think it's a really great program. I love that it combines so many different resources as far as manipulatives and technology and it has so many visuals within the book and uses so many references to everyday life." -- 4th Grade enVisionMATH teacher

In addition, ratings from both treatment and control teachers showed that enVisionMATH teachers overall were much more satisfied with their math program than control teachers. Figure 32 shows that 92.3% of treatment teachers agreed that their math program was an effective tool in math instruction compared to only 65.4% of control teachers. More than 80% of enVisionMATH teachers were satisfied overall with the program for teaching math as compared to 53.8% of control teachers. Additionally, 88.5% of enVisionMATH teachers agreed they would definitely recommend the program to other teachers as compared to only 38.5% of control teachers.

Figure 32. Teacher Attitudes of their Assigned Math Program



*Significantly different at the $p<.05$ level.

"I loved enVisionMATH! It was so teacher friendly. The students really seemed to like it better than their previous text too." 4th Grade enVisionMATH Teacher

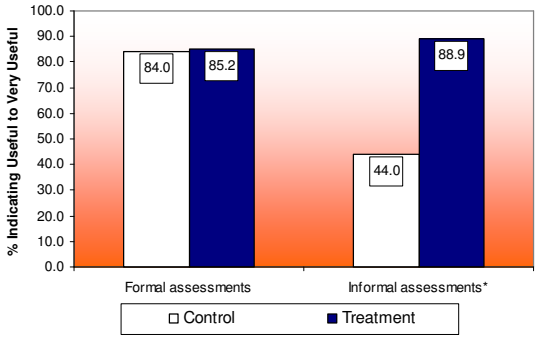
Ratings were provided by both treatment and control teachers on a variety of typical math program components. Comparison of these ratings showed significant and marginally significant differences, all in favor the of the 2009 enVisionMATH program. Specifically, enVisionMATH teachers provided more positive ratings than control teachers on the following components:

- Professional development resources embedded in the program $t(44)=4.366$, ***p*<.001**
- Informal assessments, $t(40)=4.067$, ***p*<.001**
- Review materials/exercises, $t(50)=1.928$, ***p*=.06**
- Independent practice exercises, $t(49)=2.493$, ***p* =.02**
- Technology in the program, $t(48)=3.627$, ***p*<.001**
- Manipulatives (e.g. math games), $t(49)=2.297$, ***p*=.03**
- Reading/writing in math practice provided for students, $t(50)=4.962$, p=.<001
- Program's internet resources, $t(40)=2.904$, ***p*=.01**
- Lesson planning resources, $t(48)=1.938$, ***p*=.06**
- Teaching tips, $t(48)=2.252$, ***p*=.06**

- Organization of the textbook, $t(44)=4.366$, $p<.001$
- Ease of use of the textbook, $t(50)=3.258$, $p=.10$
- Overall quality of your assigned math program, $t(50)=3.885$, $p<.001$

Overall, teachers indicated that the assessments provided with the enVisionMATH program were useful to very useful. While ratings from treatment and control teachers were similar for the formal assessments provided with the various math programs (see Figure 33), there were noticeable differences for the ratings of informal assessments. A higher percentage of enVisionMATH teachers (88.9%) indicated that the informal assessments provided with the program were useful as compared to only 44% of control teachers.

Figure 33. Teacher Attitudes about the Assessments Provided by their Math Program



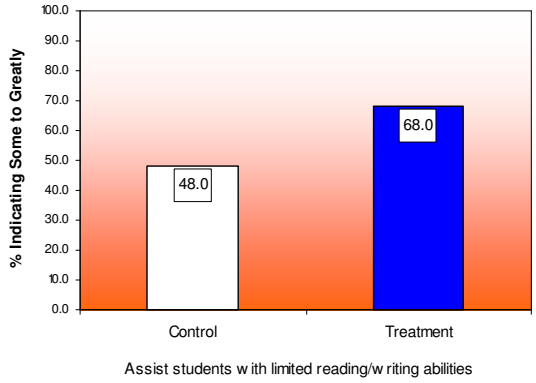
*Significantly different at the $p<.05$ level.

With regard to the remediation and enrichment resources provided by the math program, enVisionMATH teachers noted that these resources were useful. That is, teachers noted that the program was useful with differentiating instruction and reaching different types of students.

"I think it's great for all ability levels, especially with LEP and I think it's great for all students. With my low performing it gives the visuals, for my LEPs it gives those visuals plus the definition cards…With my high level they have that enrichment, (and although) I didn't get to use it that much, that's what I would give them if they mastered the lesson." - 4th Grade enVisionMATH teacher

As shown in Figure 34, 68% of treatment teachers felt that the enVisionMATH program assisted "some" to "greatly" in teaching students with limited reading/writing abilities (including students with Limited English Proficiency). Fewer than half of control teachers (48%) indicated that their programs helped address the needs of these students.

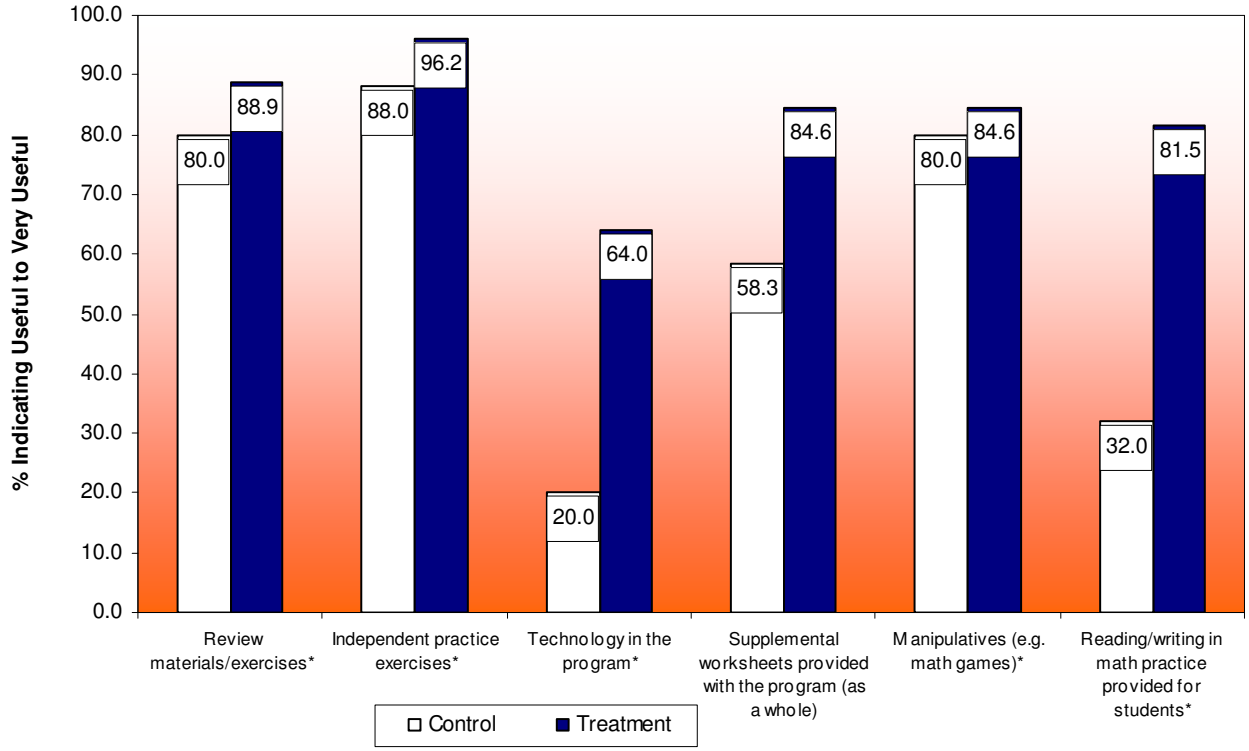
Figure 34. Teacher Perceptions of Assistance Provided by their Math Program to Limited Reading/Writing Proficiency Students



Significantly different at the $p<.05$ level.

Treatment teachers also noted that the enVisionMATH program provided sufficient additional resources to meet their instructional needs and found these resources to be useful. Moreover, ratings from treatment teachers were higher overall than ratings from control teachers, see Figure 35. Of note are the differences on teacher ratings of the reading/writing in math practice (81.5% enVisionMATH versus 32% control) and the technology

Figure 35. Teacher Attitudes about the Resources Provided by their Math Program

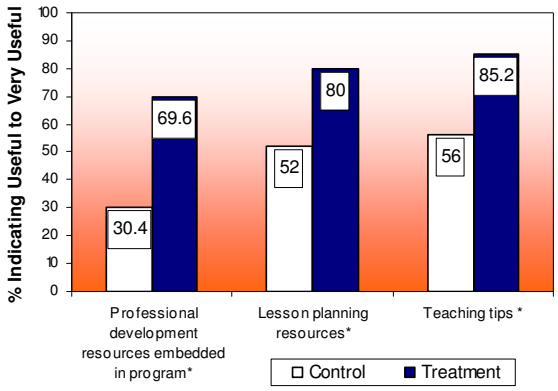


*Significant differences.

resources (64% enVisionMATH versus 20% control).

Overall, the resources provided to teachers for purposes of professional development and planning by the enVisionMATH program were rated as useful. Figure 36 shows that 69.6% of teachers using the enVisionMATH program found the professional development resources useful, 80% rated the lesson planning resources as useful, and 85.2% felt the teaching tips embedded in the program were useful, as compared to 30.4%, 52% and 56% of control teachers, respectively.

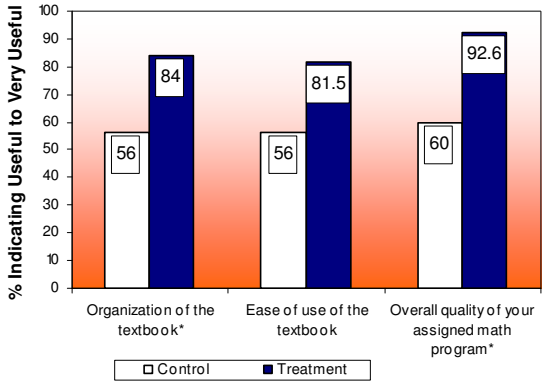
Figure 36. Teacher Attitudes about the Teaching Tools Provided by their Math Program



*Significantly different at the $p<.05$ level.

To summarize, enVisionMATH teachers liked the program overall and as shown in Figure 37, ratings of the enVisionMATH program were higher than that of control programs. enVisionMATH teachers noted that the overall quality of the program, organization, and ease of use were high. In addition, enVisionMATH teachers liked the overall layout of the program and felt that the design and graphics added increased appeal to students.

Figure 37. Teacher Attitudes about their Math Program Overall



*Significantly different at the $p<.05$ level.

"I love the layout. It's so colorful the kids enjoy it" - 2nd Grade enVisionMATH Teacher

"The way the lessons are taught is very appropriate for second graders." - 2nd Grade enVisionMATH Teacher

"I liked the organization of the topics. I really liked the flow of everything as far as how {the program} went." - 4th Grade enVisionMATH Teacher

Teachers noted a variety of specific program components when asked to identify the three things they liked best about the enVisionMATH program. However, a few items emerged as favorites from many teachers, including:

- The amount of story problems and the "thinking about math" involved in the program
- The Daily Spiral Review
- The visual representation of concepts throughout the program
- The manipulatives provided and hands-on aspects of the program
- The amount of writing provided in the program
- The design of the pouches and Teacher's Edition
- The technology and online resources provided

"I'm very, very pleased with {the program}. I like all the problem-solving compared to the previous series I used." -- 4th Grade enVisionMATH Teacher

"Having it (resources) online was a major help. I'll be at home and if I forgot my book I can pull everything I need or go over my lesson plan or if I'm out of town I can figure it out online." -- 4th Grade enVisionMATH Teacher

While overall teachers liked the enVisionMATH program, they also had some very useful feedback about the program and potential areas for improvement. This information will be provided to Pearson to inform future product revisions. The primary area that teachers noted as needing improvement was related to the pacing of the program. Teachers noted that there was a great deal of material in the lessons and that it was an adjustment to get the pacing of the program down.

"I think I'll be much more prepared to take advantage of the programs strengths next year after having used it this year". – 2nd grade enVisionMATH teacher

That said, as a two-year study, it is expected that 3rd and 5th grade teachers will have a better handle of how to implement

the program and will have a head start as compared to this year's 2nd and 4th grade teachers. Since the majority of the 2008-09 teachers have used enVisionMATH for a year, pacing issues should be less of a problem.

Teachers who used enVisionMATH overwhelmingly liked the program. A number of differences were observed in the ratings provided by teachers with respect to their math program, and these differences were always in favor of the enVisionMATH program. In other words, enVisionMATH teachers liked their math program much more than control teachers.

Conclusion

The results obtained from this randomized control trial indicate that enVisionMATH is significantly related to positive student outcomes. Elementary students using the program showed significant growth in math skills and knowledge from pre- to post-testing. Moreover, significant differences were observed between enVisionMATH and control students' performance. Students using enVisionMATH showed more improvement than control students in the areas of math computation, math vocabulary, and math problem-solving and communication.

Results also showed a number of significant differences between treatment and control students who were 4th graders, female, and minority. In particular, students in these subgroups that used enVisionMATH showed greater growth in performance from pre- to post-testing as compared to students that did not use enVisionMATH. Furthermore, students who scored above the 66th percentile during pretesting showed more accelerated gains in math vocabulary than high level math students who used other math programs.

In addition, significant differences between enVisionMATH and control students were observed across a number of different schools who used a variety of distinct curricula. Specifically, enVisionMATH students performed significantly better than control students who used programs that were purely investigative and inquiry-based as well as students who used more traditional basal math programs. This consistency in findings across different curricula, schools, and measures provides further support for the conclusion that enVisionMATH positively impacts student math performance.

Results also showed that enVisionMATH had positive effects on teacher attitudes and teacher practices. In particular, enVisionMATH teachers noted that they were more prepared to carry out various mathematics activities, and in fact, tended to engage in a greater variety of mathematics activities and strategies as compared to control teachers. This suggests that the *enVisionMATH* program has a positive impact on teacher's level of preparation to teach mathematics, and this in turn can lead to improvement in their pedagogical practices.

Teachers who used enVisionMATH also reported greater confidence to teach math as compared to control teachers and showed a significant increase in their perceived knowledge of NCTM standards and focal points.

In addition, results showed that enVisionMATH students enjoyed learning math more and perceived greater teacher

support as compared to control students. Furthermore, it was reported that students who used enVisionMATH also enjoyed the program. Indeed, enVisionMATH students rated their math program higher than control students.

The enVisionMATH program was also highly regarded by the vast majority of teachers. A full 92% of enVisionMATH teachers surveyed indicated that the program was an effective tool for mathematics instruction. In addition, 89% of enVisionMATH teachers indicated that they would recommend the program in comparison to only 39% of control teachers. Comparisons also revealed that enVisionMATH teachers rated their math programs' resources (e.g., professional development embedded within program, reading/writing in math activities, technology resources, review materials, etc) as more useful compared to control teachers.

The effect sizes for the main program effects ranged from .20-.24. While these can be classified as small effects, it should be noted that such small effects are typical of educational curricular research, particularly when comparisons are being made across similar content and classrooms, and the fact that these findings come from the first year of a two year evaluation. Indeed, positive program effects were observed despite the following:

- This was a new program for the schools and they only implemented it for one school year. Program effects take time to develop as teachers and their students become more accustomed to the program and its resources. Indeed, the number one issue brought up by participating teachers was that they had problems with the pacing of the enVisionMATH program.
- The content that was taught was similar. After all, these are not supplemental programs but core math curricula. As such, teachers, regardless of program, will be teaching multiplication, decimals, and so forth. Because of this, comparisons of two programs (enVisionMATH versus controls) which teach similar content are likely to yield small effect sizes. Typically, effect sizes for educational programs range from small to moderate (or .20-.50).

Given these caveats, results were consistent in a number of ways: (1) enVisionMATH students significantly performed better than control students on three different national assessment measures; 2) enVisionMATH students outperformed control students in different math areas (math vocabulary, math computation, and math problem-solving and communication); 3) fourteen out of the sixteen subgroup effects obtained were in favor of enVisionMATH students; and 4) enVisionMATH students showed more significant gains than control students using purely inquiry-based math programs as well as traditional basal programs, across different schools.

Activities to be Conducted During Year 2

While most of the study procedures described herein will apply during the second year of the enVisionMATH study, there are a few notable changes and additions to the methods employed during the first year. This includes:

- As a result of two schools (A and D) adopting a school-wide math curricula, these schools will not be participating during the second year of the enVisionMATH study.
- During the 2008-09 school year, data collection will be focused at the 3rd and 5th grade levels. Thus, a cohort of students will be followed over a two-year time period (from 2nd to 3rd grade and from 4th to 5th grade). Note that, with this design, the length of time a student has been exposed to enVisionMATH will vary from no exposure, to one year of exposure to two years of exposure. Hence, analyses will be conducted which specifically looks at the impact of enVisionMATH as a function of how long a student has been exposed to this program. In addition, having a group with more extended exposure to the program (e.g., two years) should contribute positively to the effect sizes observed.
- While the focus during year two will be on 3rd and 5th grade classrooms, data from 4th graders will also be collected from select schools in order to enhance the sample size, especially given the departure of two schools.
- Researchers will also obtain any existing assessment data already collected by school(s) from 2nd to 5th graders for the year(s) preceding and during the school's participation in the study. If such data is available (and accessible) it would permit cross-sectional analyses to be conducted in addition to the cohort analyses. For example, researchers can look at the performance of 3rd graders over time -- before and after the new 2009 enVisionMATH program was adopted.

In summary, the first year of the RCT enabled PRES Associates to determine that the enVisionMATH program *did* produce more significant learning gains relative to classrooms that did not use this program. Students who used enVisionMATH outperformed students that did not in the areas of math computation, math vocabulary, and math problem-solving and communication. While these main findings can be classified as small effects, if the current pattern of results hold, larger effects can be expected following two years of enVisionMATH as teachers become more familiar with the program and its resources, and treatment students have prolonged exposure to the program.

References

Cortina, J. M. & Nouri, H. (2000). *Effect size for ANOVA designs.* Sage University Papers Series on Quantitative Applications in the Social Sciences, 07-129. Thousand Oaks, CA: Sage.

DeGraw, M. (Ed.) (1999). *Balanced Assessment of Mathematics – Elementary Package 2.* Parsippany, NJ: Dale Seymour.

Gersten, R., Lloyd, J. W., & Baker, S. (1998). *Designing high quality research in special education: Group experimental designs*. Washington DC: U.S. Department of Education.

Glenn, J. (2000). *Before it's too late: A report to the nation from the National Commission on Mathematics and Science Teaching for the 21st Century*. Washington, DC: U.S. Department of Education.

Mullis, V. S., Martin, M. O., & Foy, P. (2005). *TIMSS 2003 International Report on achievement in the mathematics cognitive domains*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center. Retrieved online September 25, 2006, from http://timss.bc.edu/PDF/t03_download/T03MCOGDRPT.pdf.

National Association for the Education of Young Children (2008). *Early Childhood Mathematics: Promoting Good Beginnings.* Retrieved September 4, 2008 from http://www.naeyc.org/about/positions/psmath.asp.

National Council of Teachers of Mathematics (2005). *Mathematics Assessment Sampler: Grade 3 -5* (Ed. Gawronski, J. D.). Reston, VA: National Council of Teachers of Mathematics.

National Research Council. (2001). *Adding it up: Helping children learn mathematics*. Mathematics Learning Study Committee, J. Kilpatrick, J. Swafford, B. Findell, Editors. Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

National Science Foundation (2002). Science and Math Indicators -2002. Retrieved November 7, 2006 from http://www.nsf.gov/statistics/seind02/c1/c1s1.htm.

Rosnow, R. L., & Rosenthal, R. (1996). Computing contrasts, effect sizes, and counternulls on other people's published data: General procedures for research consumers. *Psychological Methods, 1,* 331-340.

US Department of Education (2003). *Description of Mathematics and Science Initiative***.** (Washington, DC: Author). Retrieved online February 16, 2005 from http://www.ed.gov/rschstat/research/progs/mathscience/describe.html.

US Department of Education (2006). *Math Now: Advancing Math Education in Elementary and Middle School*. Retrieved online October 18, 2006 from http://www.ed.gov/about/inits/ed/competitiveness/math-now.html.

US Department of Education (2008). *Foundations of Success; The Final Report of the National Mathematics Advisory Panel*. Retrieved September 5, 2008 from http://www.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf.

Appendix A

Crosswalk between Study Design Characteristics and WWC Review Standards

Table A1: Crosswalk between Study Design Characteristics and WWC Review Standards

WWC Causal Evidence Standards	Study Characteristics	Reference
Randomization: Were participants placed into groups randomly?	Teachers were randomly assigned to control and treatment groups within schools[32]. Random assignment was conducted via SPSS Random Selection feature by PRES researcher.	Final Report, pg. 11
Baseline Equivalence: Were the groups comparable at baseline, or was incomparability addressed by the study authors and reflected in the effect size estimate?	Randomization was reasonably successful in producing equivalent treatment and control groups in terms of student and teacher background characteristics. Still, a few differences were observed and covariates were included in the multilevel models to statistically equate the two groups and to increase the power of these analyses. These are reflected in effect size estimates.	Technical Report, pgs. 9-11 & Appendix A
Differential Attrition: Is there a differential attrition problem that is not accounted for in the analysis?	Both measurement and dropout attrition was examined. ▪ While there was no evidence for differential dropout attrition (less than 7%), there were performance differences among those who dropped out of the study. Overall, dropouts tended to have lower math scores on the BAM and MAT8 Concepts and Problem-solving subtests than non-dropouts. However, these differences were consistent between groups. As such, the threat to internal validity is minimized. ▪ Furthermore, while there was some evidence for measurement attrition between groups (i.e., a higher proportion of control students did not provide BAM posttests), results also showed that there were no notable performance differences among those who did or did not submit test performance data and group.	Technical Report, pgs. 13-15
Overall Attrition: Is there a severe overall attrition problem that is not accounted for in the analysis?	There was an overall attrition of 3.8% due to students leaving school. Note that this was part of the initial site selection criteria; in order to minimize attrition, historical mobility rates were examined and sites with high attrition rates were eliminated from consideration.	Technical Report, pgs. 14-15
Disruption: Is there evidence of a changed expectancy/ novelty/disruption, a local history event, or any other intervention contaminants?	There was no evidence of changed disruption, or a local history event. Contamination among control group teachers was also not observed. Potential treatment contaminants included: 1) the less than desirable implementation of the program by 9 of the 20 teachers, 2) the late arrival of some of the key enVisionMATH program components, and 3) the initial slow pacing of treatment teachers who were new to the enVisionMATH program.	Final Report, pgs. 15, 21, & 28-29

[32] There are a number of reasons why random assignment to treatment conditions was done at the teacher level within schools. The most important reason for selecting this level of assignment is that such a design helps to establish causality by reducing the threat that school-level factors could have potentially contributed to differences between treatment and control groups. That is, school "A" might have had something else going on (besides the treatment) that may have influenced student performance on the outcome measures. Since treatment and control groups were within the same school, school-level explanations of differences were reduced. Another reason for within school assignment is that it is likely that the treatment and control groups will possess similar characteristics at the onset of the study and therefore enhance comparability. Third, one of the criteria put forth by the DIAD study is that treatment and control groups need to be drawn from the same local pool (Valentine & Cooper, 2003). The definition of local pool provided in this study refers to subjects within the same classroom or school. According to the criteria, randomization at the district level would not be drawing people from the same local pool. Note, while this may increase the potential threat of contamination this was contained by an in-depth study orientation, monthly teacher logs, and site visits (see pages 11-13). Notably while random assignment at the teacher level within schools helps researchers control for school level differences as potential explanations of observed differences between treatment and control groups, teacher level factors can also be present and are important predictors of student performance (Gersten, Lloyd, & Baker, 1998). Though random assignment at the teacher level should help address this, with smaller sample sizes it is less likely that group equivalence will be ensured. In order to address this potential threat to initial group equivalence, additional data was collected on teacher background and classroom practices and examined and taken into account in interpretation of results. The reason why random assignment was *not* done at the lower levels (i.e., within classrooms or at the student level), was because (1) the threat of contamination if the same teacher taught both curricula was considered too great, and (2) it is not practical to randomly assign students to conditions. Indeed, schools rarely allow outside researchers to randomly assign students to use one program over another and use of the school's scheduling system as a source for randomization is not acceptable due to the fact that students' previous class schedule, ability level, student/parental requests, and so forth are factored in; this is not random.

WWC Causal Evidence Standards	Study Characteristics	Reference
Intervention Fidelity: ***1. Documentation*: Is the intervention described at a level of detail that would allow its replication by other implementers?** ***2. Fidelity*: Is there evidence that the intervention was implemented in a manner similar to the way it was defined?**	**1. Documentation:** The implementation guidelines provided in Appendix E clearly outline the expectations for implementation of the program. The enVisionMATH program is described herein in sufficient detail and references for further documentation from the publisher are provided. **2. Fidelity:** Extensive procedures were put in place to measure fidelity of intervention including training, implementation guidelines, monthly teacher logs, and site visits. Overall fidelity of implementation can be characterized as high.	1. Final report pgs. 18-20 and Appendices C and E 2. Final report pgs. 28-29
Outcome Measures: **1. *Reliability:* Is there evidence that the scores on the outcome measure were acceptably reliable?** **2. *Alignment:* Is there evidence that the outcome measure was over aligned to the intervention?**	**1. Reliability:** The assessments employed are reliable and valid. The reliability estimates for all assessments range from .68 to .96. In addition, the publishers have further information on the validity of this test. **2. Alignment**: These tests measure math concepts and skills taught in typical elementary math courses. The content reflects textbook series available from various publishers, curricula from most states, mathematics literature, and the NCTM Curriculum and Evaluation Standards and subsequent NCTM publications. In addition, the battery of assessments offer a broad coverage of content matter and consists of multiple-choice, constructed response, and computational problems.	1. Final report pgs. 12-14 2. Final report pgs. 12-14
People, Settings, and Timing: **1. *Outcome Timing*: Does the study measure the outcome at a time appropriate for capturing the intervention's effect?** **2. *Subgroup Variation*: Does the study include important variations in subgroups?** **3. *Setting Variation:* Does the study include important variations in study settings?** **4. *Outcome Variation:* Does the study include important variations in study outcomes?**	**1. Outcome Timing:** Post measures were taken within 1 month of the end of the school year. Pretest measures were taken within 6 weeks of the beginning of the school year. **2. Subgroup Variation:** The sample includes variations in gender, race/ethnicity, math ability, free/reduced lunch status, special education status, and grade level. Analyses were conducted by all subgroups, although small sample sizes among some subgroups means that results should be interpreted with caution. **3. Setting Variation:** Sites were in suburban and urban settings and in 8 states across the US. All schools were public with an enrollment (240-733 students) that is typical of schools at this level and in such settings (see Appendix B for site summaries). **4. Outcome Variation**: Four (sub)tests were used to measure the effect of the program on student performance. These included the MAT8 Concepts and Problem-solving and MAT8 Math Computation subtests, the GMADE: Math Vocabulary subtest, and the BAM. In addition, the impact of the program on student and teacher attitudes and classroom practices was also examined.	1. Final Report pg. 12 2. Final Report pgs. 24-25 3. Final Report pg. 21-23 and Appendix B 4. Final Report pgs. 12-14

WWC Causal Evidence Standards	Study Characteristics	Reference
Testing Within Subgroups: **1. *Analysis by Subgroup*: Can effects be estimated for important subgroups of participants?** **2. *Analysis by Setting*: Can effects be estimated for important variations in settings?** **3. *Analysis by Outcome Measures*: Can effects be estimated for important variations in outcomes?** **4. *Analysis by Type of Implementation*: Can effects be estimated for important variations in the intervention?**	**1. Analysis by Subgroup:** Effects were estimated via multilevel models for the subpopulations that we had sufficient data for (i.e., gender, ethnicity, free/reduced lunch status, special education status, math ability, grade level). Preliminary results showed significant program effects for 4th graders, females, minorities, and students of high math ability. **2. Analysis by Setting:** Preliminary analyses by setting consisted of examining program effects by school. These analyses showed a significant interaction with a number of schools. For the most part, positive program effects were observed (though School E showed a negative program effect). **3. Analysis by Outcome Measures:** Effects were estimated for each subtest and affective outcomes. **4. Analysis by Type of Implementation:** Program effects were estimated by variations in implementation. Results showed a significant relationship between implementation levels and math performance, such that high fidelity of implementation was associated with greater improvements in math.	1. Technical Report pgs. 21-22 2. Technical Report pgs. 21-22 3. Technical Report pgs. 19-21 4. Technical Report pg. 23
Analysis: **1. *Statistical Independence*: Are the students statistically independent or, if there is dependence, can it be addressed in the analysis?** **2. *Statistical Assumptions*: Are statistical assumptions necessary for analysis met?** **3. *Precision of Estimate*: Is the sample large enough for sufficiently precise estimates of effects?**	**1. Statistical Independence:** Analysis of the intraclass correlations showed that dependency was an issue among this sample of students. However, this was addressed by using hierarchical linear modeling and inclusion of cluster-level covariates. **2. Statistical Assumptions:** All underlying statistical assumptions were met. **3. Precision of Estimate:** Power analyses revealed that multilevel models have enough power to detect medium to large effects, as well as small effects with the use of strong covariates.	1. Technical Report pg. 7-8 2. Technical Report pg. 7-8 3. Technical Report pgs. 11-13
Reporting: **1. *Complete Reporting*: Are findings reported for most of the important measured outcomes?** **2. *Formula*: Can effects be estimated using the standard formula (or an algebraic equivalent)?**	**1. Complete Reporting**: All main findings for the outcomes are presented in the Technical Report. **2. Formula:** All effect sizes (Cohen's d) for outcomes measures are calculated and presented in the report. The formula for calculating effect sizes of main program outcomes is presented in Appendix A of the Technical Report.	1. All of Technical Report 2. Technical Report Appendix A

Appendix B:

Case Study of Site Visits

CASE STUDY OF SITE VISITS

Site visits are critical in order to better understand the context in which a program is being used. In addition, environmental factors (e.g. school factors, local history effects) can influence the results of a study and it is necessary, at the very least, to document such factors. The case study of site visits is accomplished by triangulating the data from the site/classroom observations, post-observation interviews, and the implementation logs, and capturing the perspectives of various participants[33]. The following provides information about each of the sites, collected from the participating teachers, school administrators, and our own school-related research.

School A – Colorado

About the School: School A is a large, newer school located in suburban Colorado. The school had adequate resources such as calculators, textbooks, workbooks, library books and a rich supply of technology. In 2006 the Board of Education designated School A as a "School of Technology". The school is located in a middle class residential neighborhood and serves both a rural and middle class population, with pockets of affluence. The school's student population was limited in terms of diversity, falling slightly below state averages in minority categories and slightly above in the number of White, not Hispanic students, as follows:

- White, not Hispanic: 74%
- Hispanic: 18%
- Black, not Hispanic: 4%
- American Indian/Alaskan Native 4%
- Asian/Pacific Highlander: 1%

There were approximately 16 students per full-time faculty at the school. On average, teachers had seven years of teaching experience. Colorado tests elementary students in third through sixth grade in Reading, Writing and Math. School A's performance on state standardized tests can be categorized as slightly above average as compared to state statistics, but slightly below other schools in the district. In 2005, 66% of third through sixth grade students performed at or above proficiency in math compared to the state average of 63% and district average of 73%. In 2005 the school made their "Adequate Yearly Progress" (AYP) goals as mandated by the federal No Child Left Behind Act in reading, but did not meet the AYP in math.

***Study Participants*:** There were 129 students in 6 classes at the school participating in the study, including 55 second graders and 74 fourth graders. There were three second grade teachers (one control teacher and two treatment) and three fourth grade teachers (one control and two treatment). The average class size was 22.6 students, ranging from 20 students to 26 students per class.

Second grade teachers reported that their classes contained a broad range of abilities and in general were a mix of high, low and average performing students. The exception was the control

[33] It is important to note that, when interpreting information from such qualitative data collection techniques, the data reported consist of recurrent and shared themes that emerged. That is, comments from a single individual which are not reflective of a larger proportion of respondents are not identified as a finding or "theme."

teacher who described her class as having a slightly greater number of higher performing students than the other second grade classes. Overall teachers had classes that contained one or two students with behavioral issues, but no major problems were observed.

The fourth grade teachers reported a combination of high, low and average performing students in their classes. One treatment teacher reported a mix of abilities, while the other treatment teacher reported various learner levels, leaning more towards average and low performance level students. The control teacher had the broadest scope in classroom abilities, with both higher level students and six IEP students; of those, five had behavioral problems and/or emotional issues. The teacher was able to maintain an organized class, but it was evident that she had the greatest challenge in keeping behavioral issues at a minimum and the classroom environment was slightly more chaotic than the other teachers. Overall behavioral issues were minimal to non-existent with the exception of the control teacher, as noted above.

Of note this school had teachers, across the treatment and control groups for both the second and fourth grades, that were equally competent. As well, treatment teachers expressed positive feedback in terms of using the EnVisionMATH program. To that end, implementation of the treatment program adhered almost exactly to the implementation guidelines. The school environment overall was vibrant and conducive to learning. Generally, teachers appeared to enjoy teaching and there was a positive report observed between students and teachers in almost all the classes. There were no concerns about contamination occurring whereby the control teachers were exposed to the enVisionMATH teaching approach or materials.

Math Curriculum and Resources: The math program used by the control teachers was a constructivist based program. In the 2008-2009 school year, the school will be switching to enVisionMATH. As such, this school will not be participating in study during year 2. The control program was a relatively recent version, copyright 2004. The program is organized into 6-11 Units (depending on the grade level), rather than traditional chapters. The focus of the program is very much activity-based and focused on students thinking about and exploring math. It was similar to the enVisionMATH program in terms of interactive learning opportunities; however the control program was lighter on skill practice and application as compared to enVisionMATH.

School-wide teachers used a district created computation based skill review math warm-up to supplement the control program. As well, both treatment and control teachers used Success Maker approximately 3-4 times a week for fifteen minutes on average. Overall, treatment teachers reported that they were able to feel comfortable with enVisionMATH math early on and did not report many pacing problems

***Instructional Practices and Strategies*:** Math lessons were generally taught throughout the day. On average, math was taught for 45 - 60 minutes, five days a week, with an additional ten minutes devoted to math in the mornings. A typical treatment group lesson consisted of starting with the Daily Spiral Review or Problem of the Day as a warm up activity for approximately 10 minutes. One fourth grade treatment teacher did not use the Problem of the Day at all. The warm-up activity was then followed by the teacher introducing the new topic and leading students in an Interactive Learning activity for approximately 30 minutes. During the Interactive

Learning activity teachers presented questions, gave examples, modeled problems and or made real life connections. Students were expected to answer questions, give examples and/or work problems on the board. While some teachers did so to a greater extent than others, overall teachers encouraged higher order thinking and structured questions in a way that allowed for student elaboration. .After the Interactive Learning activity teachers worked through the first few problems of the Guided Practice and then had students work the next few problems on their own, after which teachers would go over answers with the class and this took about 10 minutes. Once the Guided Practice was complete Independent Practice was assigned. During the Independent Practice teachers took the opportunity to walk around and help struggling students one on one or would pull small groups aside to offer remediation and assistance.

Control class lessons followed a similar lesson pattern which included a warm up and/or review activity for approximately 15 minutes, followed by an interactive activity with teacher-led discussion, question posing and students responding, recording data or solving problems together. It is important to note that the control teachers did not follow the interactive learning activity with any type of independent practice, rather the interactive learning activity served as the entire lesson.

The control teachers were divided in terms of homework assignment. The second grade control teacher was not observed assigning homework, nor was it noted that homework was regularly assigned during the interview. The fourth grade control teacher reported that she liked to use fact sheets as a review for her homework assignments, due to the lack of basic skill and drill provided in the control program. The treatment teachers typically assigned homework four nights a week, with a packet being handed out on Monday and turned in on Friday. The majority of the treatment teachers used the Leveled Homework and also assigned any unfinished Independent Practice as take home.

There was a strong focus and attention to math vocabulary words across the treatment and control teachers with the exception of one control teacher who did not appear to emphasize vocabulary as much as the other teachers. Use of vocabulary was not mentioned during the interview nor was it observed. However, the second grade control teacher did comment that she felt vocabulary was a particular “strength” of the her math program.

Generally, all treatment teachers used a mix of whole group instruction, small group work and independent practice, with one exception. The second grade treatment teacher reported that she generally teaches to the whole class and does not have them work in small groups as was observed. The control teachers focused on whole class instruction, with small group work as part of the interactive learning.

For the most part teachers provided differentiated instruction to their students, with the exception of the fourth grade who did report that differentiation occurred, aside from helping students during class that needed it. In general differentiation was geared towards lower level students, while higher performing students sometimes seemed bored. Teachers kept their entire class at the same point in the curriculum, with differentiation occurring via how much extra time was spent going over the lesson and working problems, rather than assigning different materials. Teachers across the study differentiated instruction by working individually or in small group

settings. Only one treatment teacher was observed giving higher performing students something additional to do. Two other teachers, one control and one treatment self reported that they offer enrichment activities to higher level students via worksheets and/or special flash cards.

Highlights: Overall both control and treatment teachers were skilled educators and successful at keeping students engaged and promoting a positive learning environment. The treatment teachers were especially pleased with the enVisionMATH program and implemented to a very high degree. Technology also contributed to a forward seeming school-wide learning environment and students and teachers alike were adept at navigating the technology provided. Smart Boards were available in each classroom and teachers utilized this resource to a high degree. Students in the control program spent slightly more time on computer and technology based learning and assessment activities than their treatment counterparts.

School B – New Hampshire

About the School – School B is a large public school located in a middle-class, suburban community in New Hampshire. The school is in an older building and houses students in grades Pre-K-4. During the 2006-2007 school year enrollment was 569, with a student-teacher ratio of 16.

Results from the 2007 statewide math indicate that 69% of third graders and 74% of fourth graders are proficient in math as compared to the statewide average of 69% and 66%, respectively. The student population is predominantly white:

- 96% White, not Hispanic
- 2% Hispanic
- 1% American Indian/Alaskan Native
- <1% Black, not Hispanic
- <1% Asian/Pacific Islander

Approximately 8% of the students were eligible for free or reduced-price lunches, with no students of Limited English Proficiency reported.

Study Participants – During the first year, 12 teachers participated in the study: 6 second grade teachers and 6 fourth grade teachers. At each grade level, three were treatment and three were control classes. The 12 classes contained approximately 268 students, with an average class size of 22, and a range of 17 to 24.

At the second grade level, teachers characterized their classes as average, with some (<5) high-performing and low-performing students. With the exception of one treatment class, all classes had 1-3 special education students. Thus, there was a mix of ability levels within each class. This was also consistent between treatment and control classes. Their classes were also noted as typical of the student population at the school.

At the fourth grade level, the majority of teachers described the composition of their students as being broad, with an average ability level. There were also 1-3 special education students in each class. However, one treatment teacher noted that approximately half of his students were lower-performing, making his class atypical in comparison to the remaining classes. During observations and in discussions with teachers, no behavioral issues were noted at either grade level. For the most part, students were actively engaged in their lessons and on-task.

All teachers seemed quite capable in their ability to teach math. For the most part, teachers did well in engaging their students, checking for understanding, asking higher-order questions, and differentiating instruction to meet the needs of their students. The school environment was also positive and conducive to student learning.

Math Curriculum and Resources – The majority of control teachers at both the second and fourth grade levels used an older, traditional basal math program, copyright year 1998, as their core math curriculum, with a couple of exceptions. One fourth grade control teacher and one second grade control teacher noted that they used the core math program for purposes of following the math topic area, but used teacher created materials on a daily basis. The core control curriculum provided visual models throughout lessons to support concepts and help bring them to life for students, which is similar to enVisionMATH. Another similarity between the treatment and control curriculum was that the program offered a variety of resources for teachers including manipulatives and teaching resources to help provide enrichment or remediation and meet individual student needs. Given the age of the control program, the availability of all of the components for each teacher may have varied. Differences between the two programs were that the control curriculum did not include any type of built-in warm up component like enVisionMATH, and enVisionMATH had a greater emphasis on math vocabulary. Also, the program consists of a more traditional chapter approach with fewer chapters with more content in each rather than the broken out topics of enVisionMATH.

In addition to the core math program, control teachers also noted that they supplemented the program with other resources they had collected over the years or created. This included worksheets for extra practice and re-teaching, math games, and manipulatives. As previously noted, one fourth grade and one second grade control teacher, almost exclusively, used outside resources for math instruction.

No district pacing guidelines were in place, and all control teachers paced their classes based on the main math program as well as state standards. Their goals were to complete all the topic areas covered by the core math program. With the exception of two teachers, use of intervention and technological resources was rare. However, this is likely due to the lack of resources they had available from the core math program.

In contrast, treatment teachers were observed following the enVisionMATH program almost exclusively. All treatment teachers adhered to the implementation guidelines almost totally. While supplementation did occur on occasion, they supplemented for purposes of providing more math practice, other math games, or to review math facts. However, whole lessons were based on the enVisionMATH program.

Instructional Practices and Strategies – Math instruction occurred throughout the day depending on the teacher. While most teachers taught math for 60 minutes, three teachers noted teaching math from 50 to 55 minutes. All classrooms were rich in resources and facilitated student-teacher and student-student interactions. Reference posters of basic math concepts were evident, as well as computers for teacher and student use. All students also had sufficient copies of math resources (e.g., student textbooks).

Typical math instruction at both the second and fourth grade level was fairly consistent across control classrooms. Teachers would begin instruction with review of homework or prior lesson, and/or doing a warm-up activity. This took approximately 5-10 minutes. Following this activity, the lesson was taught (or continuation of prior day's lesson), which at times incorporated a hands-on activity. In this respect, the control teachers' incorporation of a problem-based activity was similar to goals of enVisionMATH's Interactive Learning Activity. That said, use of hands-on activities to discover math were more evident in second grade classes and tended to be done later during the class period. One exception to this was one second grade control teacher; this teacher tended to limit the use of problem-based, hands-on activities, math games, and small groups or centers. The lesson instruction took most of the class time (approximately 20-30 minutes). Class time concluded with independent practice (10-15 minutes). For the most part, classroom activities were done in whole group or independently, with the occasional pair/small group activity.

Treatment lessons were also somewhat uniform between the second and fourth grade classrooms. enVisionMATH teachers reported using the Daily Spiral Review and Problem of the Day for review and to introduce the lesson, with one fourth grade teacher noting that these were not used due to lack of time. The daily review activities typically took 10 minutes. Teachers would begin each lesson with Interactive Learning which included stating the purpose of the lesson, problem modeling and having students share their work. These activities took approximately 15 minutes. Following the Interactive Learning activities, vocabulary was reviewed by the teachers and the Visual Learning Band was read and discussed in a whole group format. Teachers would either use overhead projectors or have their students open their math books/mats to look at the Visual Learning Band. Teachers would then lead the Guided Practice portion of the lesson. Together, these activities consisted of 20 minutes. After this, teachers would conclude the final 15 minutes with Independent Practice and the Quick Check to assess student understanding (though two fourth grade teacher noted that this was rarely used due to lack of time and because they could determine their student's level of understanding through their own observations and questioning). One area in which implementation was less than ideal was use of the various differentiation options available from the enVisionMATH program. A lot of treatment teachers rarely made use of the Leveled Homework and student Centers due to lack of time, and instead provided differentiated instruction directly to students (e.g., providing extra assistance to those students who did not seem to grasp the lesson). As well, one fourth grade treatment teacher reported that she modified lessons, substituted activities, worksheet, etc. on a regular basis. While she did not modify the structure of the program she reported that she did not tend to use the Interactive Learning and overall, did not implement to a high degree. In addition, most teachers had students work in pairs frequently on an almost daily basis, in addition to whole group and independent work.

Homework was fairly consistent between treatment and control classes. Teachers assigned homework approximately four days a week for 20 – 30 minutes, though two teachers assigned it one to two times per week. There was also a high completion rate across both classes (98%). Homework in the control classes included exercises and flashcards from the main math program or from other resources. Homework in treatment classes consisted primarily of the enVisionMATH Homework Workbook.

In terms of assessment, control teachers gave chapter and mid-chapter tests and quizzes, in addition to informal assessments (e.g., timed math facts and observations). Similarly, treatment teachers provided topic tests in both the free-response and multiple-choice format, and checkpoint quizzes.

In terms of comparability of control and enVisionMATH classrooms, with the exception of the program-based activities, classes were similar. For example, vocabulary and math computation was equally emphasized in both types of classes. In addition, both types of classes supplemented their math lessons with daily practice of math facts. However, differences were also noted. Treatment classes tended to engage in more problem-based activities including using math concepts to solve real-world problems (via Interactive Learning activities), in-depth student explanations of their math solutions, and representations and analyses of data. In addition, there were more test preparation activities in treatment classes. While this can all be attributed to the enVisionMATH program, in contrast, control teachers tended to assign more leveled practice according to their students' level of understanding.

Highlights –All teachers were quite adept at making math engaging and interesting for their students. The vast majority of classrooms, including control classes, also engaged in some form of problem-based activities. Although teachers were quite busy and in the beginning, pacing was an issue and it was at times hectic as they were in their second year of a new language arts program, the treatment teachers quickly learned the enVisionMATH program and did well in following the implementation guidelines, with the exception of use of the differentiation resources. However, this was due to a lack of time as opposed to not liking or wanting to use the activities. Of note, is that one fourth grade control teacher left during late November which caused some stress among the students. However, a new teacher was later hired (after having a sub for a few weeks) and the students quickly recovered. There were also no notable differences between treatment and control students in terms of mathematics ability and no evidence of contamination.

School C – Massachusetts

About the School – Elementary School C is located in a middle to upper-class, suburban community in Massachusetts. Located within a residential area, it consists of a new, modern building. The building itself is divided into two schools, one consisting of grades PreK-3 and the other consisting of grades three through 5. Enrollment for the lower primary school was 733 in 2007, while enrollment at the upper primary school was 618. Although a large school, the school environment was very positive with strong educational support from the administration, paraprofessionals and other staff, and teachers. The school was well-organized and had walls

filled with student work. It was clear during site visits that staff took great pride in their students and aimed to provide a positive learning environment.

In 2006-2007, Massachusetts used the Massachusetts Comprehensive Assessment System (MCAS) to test students in third through eighth grades and tenth English language arts and math. Results from the MCAS math assessment indicate that 77% of third graders, 67% of fourth graders, and 71% of fifth graders were proficient in math according to Massachusetts state standards, as compared to the statewide average of 60%, 48%, and 51% respectively.

Approximately 2% of the students were eligible for free or reduced-price lunches, 0.5% were English Language Learners, and 16.5% were students with disabilities. In terms of ethnic/racial background, the student population was predominantly white:

- 93% White
- 2% Hispanic
- 2% Asian/Pacific Islander
- <1% Black
- 1% Multi-racial

Study Participants – During the first year, 14 teachers participated in the study: 7 second grade teachers and 7 fourth grade teachers[34]. At the second grade level, 3 were treatment and 4 were control classes. At the fourth grade level, there were 3 treatment classes and 5 control classes (note that one control teacher had two French immersion classes). The 14 classes contained approximately 339 students, with an average class size of 22, and a range of 22 to 24.

Second grade classrooms were characterized as mostly average to high performing. In all classes, teachers noted that the minority (25% or less) were low-performing students. In addition, each class had on average three to four special education students. To assist these students, there was a paraprofessional who assisted teachers in various classrooms as needed. This mix of students was consistent across treatment and control second grade classrooms.

At the fourth grade level, teachers were divided into teams according to the math program used. With the exception of the French Immersion control teacher, three control teachers rotated from one group of students to another, initially according to student performance on a prior chapter test. For example, during one month, a teacher may be teaching the lower level students and in a subsequent month, the same teacher may be teaching the advanced students. Thus, there was rotation among teachers, students, and even the classroom where instruction took place. It is also important to note that while initially, groups (or classes) where grouped by performance on tests, later in the school year they tried to keep the groups more heterogeneously grouped. Thus, 4th grade control students had a broad-range of ability levels, from low to high-performing students.

Among the fourth grade treatment classes, student performance levels were average, with a small number of high and low performing students in each class, with one exception. The majority of

[34] Of note is that 2 teachers who were initially assigned to be a part of the treatment group were subsequently dropped from the study due to extremely low implementation of the enVisionMATH program. These teachers supplemented enVisionMATH so heavily that it no longer resembled the enVisionMATH program. As a result, the teachers were excluded from the study late in the Fall.

one teacher's class was made up of lower performing students. Treatment classes overall contained a few special education students and support teachers were available on most days.

Overall, the classroom environment was positive. Students were intellectually engaged with important ideas in the lessons observed, lessons reflected careful planning and organization, there was a conversational tone to interactions in the classroom and teachers encouraged active participation of all students, and the teachers were able to focus on student learning as opposed to controlling behavioral issues.

Math Curriculum and Resources – The control teachers at both the second and fourth grade levels relied on a basal math program. Second grade classrooms used a newer version of the program (2005) whereas fourth grade teachers used an older version (1998). While the second grade control teachers and the French Immersion fourth grade control teacher relied on this math program for mathematics instruction, the remaining fourth grade control teachers who were in the rotating team noted that they used the program more as a content resource to identify what math topics to cover but supplemented heavily with other resources to actually teach the lesson.

There were a number of similarities between the control program and enVisionMATH. Similarities include the warm-up options provided in the program (Problem of the Day, Spiral Review), the Diagnosis and Intervention System and some of the manipulative and differentiation resources were also similar. However, there are also a number of differences between the two programs. Specifically, enVisionMATH includes a great deal more emphasis on developing student's math skills by increasing visual/verbal connections. The Visual Learning and Interactive investigations based activities, Center Activities and the structure of the program into new learning strands and 20 Topics that are incorporated in the enVisionMATH program are a few of the biggest differences between the two programs.

In addition to the core math program, control teachers, especially the fourth grade control teachers in the rotating team, supplemented with other resources they had collected over the years. This included Exemplars and TOPS to reinforce problem-solving strategies, worksheets, math games, "Challenge Packets," Versatiles, and manipulatives (e.g., tangrams). For the most part, these outside resources were used to practice math facts, for problem-solving practice, enrichment, and reteaching.

There were no school-wide math initiatives in place. In terms of pacing, control teachers followed district and state standards and pacing guides which aligned closely with the core math program. In contrast, treatment teachers tended to follow the enVisionMATH topics sequentially, with a few exceptions. For example, some teachers began the school year with the enVisionMATH Data Topic because a review of this math topic was needed for science. Two fourth grade teachers noted that they covered geometry concepts early on.

Treatment teachers at school C were observed following the enVisionMATH program closely and did well in following the implementation guidelines. While supplementation did occur to meet the needs of their students (e.g., for extra practice or enrichment), it is important to note that this occurred in addition to (as opposed to replacing) enVisionMATH resources. Treatment teachers supplemented with TOPS, My Skills Tutor, other teacher-created materials and games,

and use of additional manipulatives (e.g., decimal squares, polygons, Versatiles) and was consistent with use of supplementals in control classrooms.

Instructional Practices and Strategies – Math instruction occurred throughout the day depending on the teacher and day of the week (e.g., one teacher taught during the mornings two days per week and in the afternoons during the remaining days). While most teachers taught math for 60 minutes, four teachers noted teaching math from 30 to 90 minutes, depending on the day of the week. The classrooms had many resources available, including reading books, reference posters, and visuals to help reinforce important educational concepts from a variety of subject areas. Most classes also had computers available for both teacher and student use.

Typical math instruction in second grade control classes began with a hands-on activity or math game with students working with manipulatives (math games were common in two of the four control classes) or with a math review, followed by the teacher conducting whole group instruction on the math topic and then independent practice. Throughout the lesson, teachers would check for student understanding by asking students questions and/or through observation. Review of previously taught concepts was also evident during observations. Small group or center work occurred with some regularity in two of the four control classes, for the purpose of providing differentiated instruction. One teacher also noted placing great emphasis on differentiated independent practice through the use of enrichment, remediation and on-level practice worksheets. In comparison to treatment classes, control teachers rarely engaged in problem-based interactive activities.

At the fourth grade level, there was more variation in teaching styles and strategies employed though there were also some common elements. Teachers usually began the lesson with a review of vocabulary math words. A review of homework or previously math topics also occurred. Sometimes, this entailed a game (e.g., throwing a ball around once a student said the correct answer to a multiplication problem). Whole group instruction would then occur followed by independent practice when time remained. One teacher was noted as providing many more opportunities for center work in comparison to the other three teachers. Like the 2nd grade control teachers, teachers would check for student understanding by asking students questions and/or through observation. One teacher also made it a regular practice to provide differentiated independent practice or to use small groups for this purpose. It is also important to note that because students rotated in and out of different math classes every month or so, the content taught across three of the four math classes (with the exception of the French Immersion 4th grade classroom) was very similar.

Because of the consistency in format of the enVisionMATH program, treatment lessons were more uniform between the second and fourth grade classrooms. With the exception of two treatment teachers, teachers would generally begin their lessons with Daily Spiral Review and/or Problem of the Day. The two teachers noted that they rarely used these due to lack of time. As well, the Daily Review typically took 7-10 minutes. This was followed by the Interactive Learning Activity which took about 20 minutes. The Visual Learning Band was reviewed with students following along in their textbooks. This portion of the lesson typically took 7-10 minutes. The final component was the guided practice and independent practice. It should be noted that the Quick Check was not used with much frequency by most teachers; teachers noted

using this, on average, 25% of the time because they rarely had time and because they felt they could determine their student's level of understanding without this aid. Treatment teachers also did well in working with students who did not demonstrate an understanding of the lesson and to provide differentiated practice to on-level and advanced students through the various resources available with the enVisionMATH program. Similar to control classes, classroom instruction occurred in whole group primarily, but teachers also used pairs or small groups for Interactive Learning Activities or other activities.

Homework was consistent between treatment and control classes. Teachers assigned homework approximately four days a week for 20-25 minutes. There was also a high completion rate across both classes (96.5%). Homework in the control classes were oftentimes pulled from the practice workbook that came with the main math program, though they would occasionally give homework from other resources (e.g., for problem-solving or enrichment). Homework in treatment classes were obtained from various resources included the practice and enrichment worksheets, daily spiral review, and homework workbook. Control classes would typically assess their students with Chapter Tests and checkpoint quizzes. For the treatment classes, formal assessment occurred via the Topic Tests (free response and performance assessments). All second grade classes were also required to take the "School Math Skills Test" for report-card purposes.

Comparison of control and enVisionMATH classrooms revealed that the classes are similar in many respects, except for program-based activities and approaches. While vocabulary, math computation, problem-solving, and mental math were of high importance and emphasized across the majority of classrooms, enVisionMATH teachers tended to put more emphasis on test-taking skills and helping students discover alternative solutions to math problems as compared to control teachers. enVisionMATH classrooms were also more apt to engage in problem-based, interactive learning activities as compared to control classrooms and to respond to textbook questions (via the guided practice). One second grade control teacher placed a heavy emphasis on problem-solving, devoting approximately 30 minutes daily of activities designed to improve upon this math skill. As previously noted, all teachers (treatment and control) would supplement on occasion with TOPS, My Skills Tutor, and other teacher-created materials and games.

Highlights – Of note is that treatment teachers did not begin implementation of the enVisionMATH program until approximately one month following the school year. This was due to prescheduled activities that interfered with training scheduling. Overall, the teachers did well in making math fun and engaging their students in their lessons. The class and school climate was positive and encouraged a collaborative approach to teaching and a supportive environment for students. Other than program-related activities, there were no notable differences between treatment and control classes. Additionally, there was no evidence of contamination between treatment and control teachers.

School D – North Carolina

About the School – School D is a public school located in a rural area of North Carolina. It is a newer two-story facility with large classrooms. This was only the second year that the facility had been in use. Enrollment for the site was 298 students in grades PK through 5. School level demographic information is noted below. As shown, the majority of students participating in the study from School D are predominantly Caucasian (80%), followed by Hispanics (20%). At one point, this school was designated for all ESL students in the county; therefore, there is a large population of Hispanic students in comparison to other schools in the district. School D has approximately 15% of students in special education and has a large percentage of students eligible for free or reduced lunch (79%).

- 80% White
- 18% Hispanic
- 1% Native American
- <1% Black
- <1% Asian/Pacific Islander

In terms of overall school level math performance on the North Carolina End of Grade (EOG) Test, in 2006, 52% of third grade students performed at or above proficiency in math compared to the state average of 68%, and 54% of fourth grade students performed at or above proficiency compared to the 65% state average. Fifth grade students fell into the lowest percentile, with only 37% of students performing at or above proficiency compared to the state average of 63%. For all three grade levels, a much higher percentage of students performed at or above proficiency in reading (average 78%), although were still slightly lower in comparison to the state average.

Study Participants – During the 2007-08 school year, 87 second and fourth grade students (40 control; 47 treatment) in 6 classes (2 control and 1 treatment at 4th grade, and 1 control and 2 treatment at 2nd grade) participated in the first year of the RCT at School D. The average class size was 16 students, with a range of 15-19 students.

Fourth grade students were rotated out of homerooms for math, and this was the first year that they were grouped by math abilities based on their EOG scores. Because of this, there was one math class with low-performing and inclusion students (control), one for average students (enVisionMATH), and another for high-performing students (control). Second grade students were with the same teacher all day, and were placed randomly rather than based on ability. Most 2nd grade classes were described as being highly typical with a broad range of abilities represented, and the majority of students being average. The majority of 2nd classes also seemed to have at least some lower-performing and remedial students. No significant behavior issues were observed, and students generally appeared to be engaged in the lessons.

Math Curriculum and Resources – The primary math curriculum used by control teachers at School D was a traditional basal math program (2004). This curriculum is aligned to North Carolina State Standards; it aims to provide coverage of state and national standards while providing teachers flexibility to customize the program according to their state and local needs.

The control program is similar to enVisionMATH in that mathematics topics are covered in a larger number of short chapters, as compared to other traditional textbooks. Another significant similarity between the two programs is that both enVisionMATH and the control program place a great emphasis on building conceptual understanding, problem-solving and reasoning skills. Both of the programs also share the goal of carefully developing concepts within and across mathematics strands in order to increase long-term understanding among students.

In general, all control teachers had access to the complete components of the control program. The 4th grade control teachers used the control curriculum as their main resource for instruction, although they did some supplementing of the text with investigative and hand-on activities. The 2nd grade control teacher used it primarily as a pacing and content guide. However, he supplemented heavily with teacher created investigative activities and worksheets from other publishers.

The two biggest differences overall between the control program used and enVisionMATH were: 1) the control text was published as a state-specific program to match North Carolina state standards more closely; and 2) enVisionMATH incorporates an integrated approach to instruction by blending aspects of both investigative and traditional basal instruction whereas the control program focuses more on traditional instruction.

enVisionMATH teachers completed most of the key components of the enVisionMATH program and in general, implemented the program with moderate fidelity. It should also be noted that teachers did not regularly use program components such as the Diagnosis and Intervention System, and the Center Activity Kits as a result of time constraints. In addition, pacing issues posed a challenge early on in the school year. This was particularly evident among one 2nd grade classroom. However, as with any new program, teachers had to adjust, and by mid-year most concerns with regard to pacing had been resolved. Indeed, pacing increased as teachers became more familiar with the program, students had more of the required background knowledge, and teachers had additional professional development via a follow-up training session.

Instructional Practices and Strategies – Math lessons were generally taught during the first half of the school day, and were completed before noon. On average, math was taught for 60 minutes, five days per week. One 2nd grade treatment class was set up slightly differently than the others, and was broken up into two halves with lunch in between.

Sufficient resources were evident for classrooms in this school. Classrooms had 2-5 computers available for student use. While most teachers did not appear to have a computer specific for their own use, the 4th grade treatment teacher had access to her own computer and Smartboard. Projectors, a television and VCR, as well as multiple manipulatives were present in every classroom. Classrooms were all very spacious and the design facilitated center activities and game stations during the math classes.

Although some variation existed in how treatment teachers implemented their lesson, a typical enVisionMATH lesson can be described. Treatment teachers typically began their math lessons

with the Daily Spiral Review for about 10 minutes. Following this activity, teachers would introduce the next topic and review the necessary Vocabulary, before delving into the Interactive Learning activity for anywhere from 10 to 20 minutes. The Visual Learning Band was discussed and the Guided Practice would then be explicitly modeled. At this point, 2nd grade teachers would sometimes read the related section from the Big Story Book to the class before answering questions together as a class. Independent Practice usually followed for the remainder of the class period. Teachers would also have students engage in Quick Check (or other error intervention activities) and at times, Learning Centers in order to check understanding and differentiate instruction. EnVisionMATH teachers also encouraged collaborative work, sharing of ideas, making connections between the math lesson and their own experiences, and higher order thinking skills. Treatment teachers using the enVisionMATH program rarely supplemented, if any, from outside resources.

Of note is that initially, pacing was an issue for treatment teachers, as most felt that they were behind the other teachers from the same grade level. This was particularly true of one 2nd grade treatment teacher. However, at the time of the spring site visit, all the treatment teachers reported that they had significantly reduced issues with pacing. Furthermore, examination of teacher logs and implementation checklists reveals little difference in the amount of math lessons covered among treatment and control teachers.

Similar to the enVisionMATH classrooms, control teachers generally used investigative and collaborative hands-on activities in their lessons. Manipulatives were used in every classroom. Although each teacher lectured for a short time during each lesson, the majority of classes were spent in pairs and small groups. A typical 4th grade lesson consisted of review, followed by an introduction to the lesson, guided practice, and a hands-on activity. Like the enVisionMATH classrooms, control students were often engaged during their math lessons. Of note is that the 4th grade inclusion class control teacher focused on meeting the needs of her nine special needs students; the teacher emphasized the personal attention necessary to reach these students, and spent the majority of class time assisting students with investigative activities and emphasizing key vocabulary words. It should be noted that this teacher also tutors several enVisionMATH students as a group but does not use enVisionMATH in her control classroom.

All the 2nd grade classes had a teaching assistant, which they worked closely with. It was reported that the assistants often were involved in a teaching role, but overall during the observations, the assistants seemed to focus mostly on the students requiring additional help and attention, thereby allowing the teacher to focus on keeping the lesson moving. The 4th grade teachers reported that they initially had interns and/or mentors in their classes 1-2 times per week to help struggling students keep up at the beginning of the year.

Highlights – It is important to make note that initially, the treatment teachers had a difficult time with enVisionMATH in terms of pacing, supplementation, and collaboration with fellow teachers. However, during the spring site visit it was observed that the teachers felt comfortable with the pacing and had become comfortable with the program. For the most part, the overall pedagogical approaches employed by School D teachers were fairly comparable since the majority of teachers instructed using more of an inquiry-based rather than a traditional textbook approach, and used lessons that included a variety of teaching approaches and activities. In

addition, as a result of changes in administration, some teachers in the 3rd and 5th grade not interested in participating in the study, and the school's desire to use a single math curriculum in 2008-09, this school will not be participating during the second year of the study.

School E – Tennessee

About the School: School E is a small public school with approximately 249 students located in rural Tennessee. The school, which serves grades PK-5, is situated in a middle class residential area. The facility is 50 years old with one new wing built in 2005 to house the administrative offices, bathrooms, and a cafeteria. While the facility is older, it is in remarkably good shape and feels like a much newer building. There is a strong feeling of community involvement at the school and it was reported that a number of retirees mentor struggling students at least once per week. In general there is a cohesive feel to the school. Students seem to receive a high level of home support and this carries over to the teachers, who overall were very positive in terms of their students, the school and teaching in general. The support and collaboration among teachers and the administrative staff is very apparent and certainly contributes to the positive atmosphere.

The school's student population was limited in terms of diversity, as follows:

- White, not Hispanic: 96%
- Hispanic: 2%
- Black, not Hispanic: <1%
- American Indian/Alaskan Native <1%
- Asian/Pacific Highlander: <1%

Approximately 24% of students were eligible for free or reduced price lunch. There were 14 students per full-time teacher at the school. In 2006-2007 Tennessee used the Tennessee Comprehensive Assessment Program (TCAP) Achievement Test to assess students in grades 3 through 8. Of students at School E, 98% of third and fourth grade students performed at or above proficiency in math compared to the state average of 89%.

Study Participants: There were 75 students in 4 classes at the school participating in the study, including 33 second graders and 42 fourth graders. There were two second grade teachers and two fourth grade teachers, one each for treatment and control. The average class size was 19.25 students, ranging from 17 students to 22 students per class.

Participating teachers described their math classes as being average to above average overall, with some variation between control and treatment classes. Specifically, while the second grade control teacher noted that the majority of her math class was comprised of higher level students, the second grade treatment teacher described her class as being average, with students who required a slightly slower pace than the other second grade class. Similarly, the fourth grade control class was also comprised of students representing a high-range of math ability and were characterized by the teacher as "high achievers;" in contrast, the fourth grade treatment teacher described her class as being average, and required a slower pace and more individual attention than the other fourth grade class. That being said, both control and treatment classes each had a

small number of lower-performing students as part of their overall student profile. No significant behavior issues were observed for either of the control classes and students generally appeared to be engaged in the lesson.

Math Curriculum and Resources: There were no school-wide initiatives or special programs targeted at math instruction. Both control teachers primarily used a basal math program that was specific to Tennessee's state standards. The control curriculum used at the school was a basal program published in 2005. Similarities to the enVisionMATH program included mathematics topics that were covered in a larger number of short chapters, as compared to other traditional textbooks. Another significant similarity between the two programs was that both the enVisionMATH and the control program placed a great emphasis on building conceptual understanding, problem-solving and reasoning skills. Both of the programs also shared the goal of carefully developing concepts within and across mathematics strands in order to increase long-term understanding among students. The primary difference between the two programs was that the control program was designed specifically to help teachers teach to Tennessee specific standards. Additionally, enVisionMATH incorporates a more integrated approach to instruction through blending aspects of investigative and traditional basal instruction, whereas the control program focused more on traditional instruction.

In addition to primarily relying on this text, the fourth grade control teacher also supplemented with an older version of the control text and commented that she felt the control program was "light" on problem-solving and skills practice. In general, the control teachers taught using a basal instructional approach in comparison to the enVisionMATH program. Due to the fact that each grade in the school followed a fairly rigid curriculum map, most classes were on the same pace based on these guidelines. The state specific text chosen for the teachers by the county paralleled the curriculum mapping and also served as good preparation for standardized testing. On average, classroom resources were adequate. Each classroom had three computers available for student use, as well as a television and VCR.

Instructional Practices and Strategies: Math lessons were taught during the first part of the school day, though ranging in start times, five days per week. Math lessons ranged from approximately 45 to 60 minutes in length. Contamination did not appear to be an issue, as control teachers were locked into set curriculum mapping and the associated Tennessee version math text.

For control teachers' lessons generally began with a brief review, followed by the lesson which was largely lecture-style, guided practice, and independent practice near the end. As previously noted, students worked primarily out of their textbooks, with very little, if any, hands-on or investigative-type activities, with the exception of the fourth grade control teacher's Friday math games, as noted below. Both the control teachers reported that they gave assigned in-class independent work daily. The fourth grade control teacher assigned homework about once a week, while the second grade teacher reported that she rarely assigned homework, but did expect students to take unfinished work home and return with it completed on Monday. The fourth grade control teacher mentioned that she would occasionally partner students to work together if they were struggling and that she did a whole class game activity like "Math Jeopardy" teams every Friday. "Math Jeopardy" was observed and it was evident that the game required students

to use math vocabulary, mental math and to show their work under pressure. The second grade control teacher stressed that she did not put students into groups very often, nor was it observed. For both the second grade and fourth grade control math classes, instruction was most often led by the teacher with little interaction between the students, with the exceptions as noted above. Both teachers did ask for student participation in the beginning of the lectures and had students volunteer to come to the front of the room to solve problems on the board or explain how they arrived at their answer.

Overall the treatment teachers conducted lessons in a similar format with a few exceptions. The second grade treatment teacher reported that she did not use the Problem of the Day, Daily Spiral Review or Quick Check at all. Thus, while the 2nd grade treatment teacher implemented the key components of the program to a lower degree, the 4th grade treatment teacher showed a high degree of implementation fidelity. As an example of a lesson observed, the teacher started with a review of the previous day's lesson on decimals which lasted approximately 10 minutes. The teacher asked students to come up with real life examples of when they had seen or encountered fractions and explain how and why the decimals were used. Students were very engaged and applied higher order thinking in coming up with answers. The teacher then clearly stated the lesson objective for the day, which was multiplying with decimals. She did an activity where she had students plan a pizza party and used the slices of pizza as the context for multiplying fractions (note that this activity was used in place of the enVisionMATH Interactive Learning activity), which took approximately 20 minutes. She used the Visual Learning Band in the book to develop the concept. This was followed by Guided Practice and then Independent Practice, which took 10 minutes and 15 minutes respectively. As part of the Independent Practice, the fourth grade treatment teacher assigned the Quick Check problems and used this to gauge student comprehension and followed this with re-teaching. The fourth grade teacher also reported that she did the Daily Review or Problem of the Day in the morning. It was noted by both teachers that they did not use the Center Activities. The fourth grade treatment teacher reported that she tried the Center Activities early on, but stopped using them altogether, as she found it too distracting for her class, while the second grade treatment teacher reported that they were too time consuming.

It should be noted that the fourth grade treatment teacher felt that her students had a hard time getting used to the format of the enVisionMATH program at the beginning of the year, as it required them to think in ways they had not been exposed to in earlier grades. Specifically, her students had a hard time conceptualizing that there was more than one way to solve a problem for the correct answer. She felt that towards the end of the year they had really picked up on this concept and were better able to apply this approach.

In general, treatment and control classes were similar, with a few notable exceptions. For the most part, while the treatment teachers engaged in a slightly more investigative *style* than the control teachers, especially the fourth grade treatment teacher. Overall the teachers were more traditional in their pedagogy. All of the math classes tended to be taught in a whole class format and were extremely structured, with the exception of the second grade treatment teacher's class who had a more unstructured classroom environment. In addition, across the classes teachers tended to do little if any differentiation with the exception of the second grade control teacher who did assign challenge work to her more advanced students. All the teachers required that

their students show their work and explain how they arrived at answers with the exception of the second grade control teacher. In addition, all of the teachers did a timed math facts test everyday, though it did not always occur during the actual math period.

In terms of math content, as a result of the curriculum map and pacing guidelines, treatment and control teachers taught similar math topic areas (e.g., decimals, multiplication, etc.). In addition, vocabulary was a focus across the classes with the exception of the second grade control teacher, who did not seem to emphasize this as much. Due to the nature of the enVisionMATH program, and the apparent lack of problem-solving and skill practice in the control program, treatment teachers were also more likely to expose their students to problem-solving skills.

In addition to the slightly more investigative-style and problem-based teaching that the treatment teachers engaged in, the other notable difference between treatment and control classes were the student ability levels within these classes. As previously noted, both of the treatment classes were the lower performing class at each grade level, while the control classes were primarily comprised of the more advanced students.

Highlights: It is important to note that School E places a large focus on literacy in terms of their overall curriculum. It was evident that reading, vocabulary, and language were a top priority at this school. In all of the observed classrooms the walls were almost entirely devoted to literacy related content, with very few math related posters or charts. Overall the school was very organized and there seemed to be an unspoken structure that governed everything. Of note is that the second grade treatment teacher had to deal with a series of family issues that affected her attendance throughout the year. In addition, while it was apparent that teachers got together frequently to discuss students and to support one another, no signs of contamination were apparent.

School F – Kentucky

About the School – School F is located in rural northwest Kentucky, in an area dominated by farmland. The school building itself was built only recently (2005), housing grades K-5 in a modern facility. Enrollment during the 2006-2007 school year was 514, with a student-teacher ratio of 18.

School F used the Kentucky Core Content Tests (KCCT) to assess students in 2006-2007. Results from last year's test indicate that students in grades three through five produced a combined scored of 99 on the academic index, with KCCT indices ranging from 0 to 140. The statewide goal for all students was 100, although the statewide average for math was 90.

The student population is predominantly white:

- 97% White;
- 2% Hispanic
- <1% Black;
- <1% Asian/Pacific Islander.

Approximately 40% of the students were eligible for free or reduced-price lunches, with no students of Limited English Proficiency reported. School-wide, approximately 9.2% of students are in special education.

Study Participants – During the first year, six teachers participated in the study: four second grade teachers and two fourth grade teachers. Of the second grade teachers, two were treatment and two were control classes. At the fourth grade level, each teacher is responsible for two classes; therefore, there are two treatment classes and two control classes at fourth grade. The eight classes were composed of approximately 174 students, an average class size of 22, with a range of 17 to 24.

Fourth grade teachers each described the composition of students in their classes as fairly typical in terms of ability, although both recognized that the classes were tiered. The control and treatment teachers characterized their first period math classes as low- to average-range and their second period math classes as high-range. All four second grade teachers considered their students average in comparison to students in other second grade classes, and described their class' ability as either mid-range or broad-range. No behavioral issues of any kind were observed at either grade level.

Math Curriculum and Resources – Control teachers at both the second and fourth grade levels used their assigned core control program almost exclusively for math instruction. The control curriculum used by teachers at the school was a 2007 copyright (3rd edition) of basal math program. The design of the program is intended to develop concepts and skills over time in a variety of contexts as well as to teach students there are multiple methods and strategies for problem-solving. The program is similar to enVisionMATH in two main ways; 1) each lesson begins with a very focused "objective" to help guide instruction and provide students with a sense of the goal for what they should be learning; and 2) a focus on differentiation options for supporting the needs of all learners. The biggest difference between the two programs is that the control program does not come with manipulatives or include as much emphasis on investigation/hands-on activities.

No district pacing guidelines were in place, and all three control teachers paced their classes based on the main math program. The fourth grade control teacher noted that the math program lacked sufficient drill and practice fact problems for students at that level, and supplemented with a timed practice review every Friday. Both second grade control teachers were observed using more hands-on manipulatives with their students, from Geoboards to counters, and commented that they supplement with flashcards from an outside the program and occasionally pull worksheets from the computerized Success Maker program. Additionally, second grade control teachers also supplemented with the use of centers and computer programs such as Digitar to practice math facts.

In contrast, treatment teachers at Elementary School F were observed following the enVisionMATH program almost exclusively. Specifically, both second grade treatment teachers adhered to the implementation guidelines almost totally and rarely supplemented with outside

materials. The fourth grade treatment teacher also adhered to the program, though supplementation occurred occasionally through use of materials she had collected over the years.

It is important to note that both treatment and control fourth grade teachers began the year with a self-developed geometry unit that was not based on their respective programs. As a result, implementation of the enVisionMATH treatment program did not begin until several weeks into the school year. Additionally, while used as a supplement for control teachers, Success Maker was also employed by treatment teachers as part of a school-wide initiative to target math skills. Students used the individualized computer program for approximately 30 minutes a day, spending half the time on reading skills and the remaining half on math skills.

Instructional Practices and Strategies – Math was instructed in the mornings at the start of the school day in fourth grade classrooms, with both the treatment and the control teacher instructing first and second periods for approximately 55 minutes, from 9am until 11am. Math instruction at the second grade level varied slightly more, ranging anywhere from 50 to 70 minutes. Both second grade treatment teachers reported spending an average of 60 minutes each day instructing math with the enVisionMATH program. All classrooms at Elementary School F were above average in terms of resources, with new, modern tools to facilitate student learning including computers for both teacher and student use. Most classrooms were overly spacious and rich in materials for student use.

Typical math instruction at the second grade level varied slightly between control teachers, but was usually student-centered and hands-on. In the control classroom that spent on average 70 minutes a day on math instruction, the teacher would introduce the day's concept with a kinesthetic activity to engage student interest. After a brief, 5 minute introduction to the hands-on task, students would spend approximately 30 minutes working with the investigative activity either individually or in small groups as the teacher monitored student progress. Following this task, the teacher would redirect students to an activity or worksheet from the student text, and students would engage in independent practice at their desks. During this time and for the remaining 25 minutes, the teacher would work with students requiring assistance in small groups or walk around to help students individually as needed. Time permitting, the class would engage in flashcard activities at the conclusion of class to wrap-up. This teacher reported assigning math homework very infrequently – once or twice a grading period – and using teacher observation, lesson reviews, and occasional workbook pages as the major forms of student assessment, as unit tests were few.

In the remaining second grade control classroom with instruction averaging around 50-60 minutes daily, the teacher would set up the day's math lesson with a Morning Math Message or Math Journal Topic at the start of the school day and return to it at the beginning of the math period. Following this brief 10-15 minute review, the teacher would recap the previous math lesson or simply progress straight on the Interwrite (a computer presentation system for displaying information and manipulating objects on a projector screen) which typically consisted of modeling several problems and math processes on the projector screen. This would lead directly into an investigative activity or students' workbooks for independent practice. During the remaining 10 minutes or so that students were allowed to work in class, the teacher would encourage the use of manipulatives for student assistance, and walk around to students' desks to

help those requiring personal attention. The following day's instruction would consist of a spiral review workbook page from the text, and the subsequent day would typically involve an enhancement activity or review session if reteaching proved necessary. Homework was sent home on Mondays and would consist of 2-3 assignments averaging 10-20 minutes in length. Homework would then be checked on Fridays as part of the "catch-up" day of the week; other forms of assessment included daily worksheets completed as a whole group as well as the end-of-chapter tests given. Occasionally, open-response assessments would be given to the entire class.

Instruction in the fourth grade control classroom would usually begin with a "Math Message" at the start of the lesson as part of the program and a follow-up discussion as a readiness activity. Spending about 20-25 minutes on this review, the teacher would then model several problems on the Interwrite system for the class before distributing a worksheet for students to complete in small groups with the aid of manipulatives while the teacher and assistant circulated to help students. Small group activities were also employed to reinforce the math being taught. The remaining class time (usually 10 minutes) would be spent working on a spiral review page or flashcard activities to reinforce math facts. Once a week, students would complete a writing-to-explain assignment, and every Friday the teacher would assign a timed practice skills test. Homework was assigned as part of an online "Study Links" website every week on Monday and was due the following Thursday, with assignments averaging between 30 and 45 minutes. In addition to the weekly timed tests and open-response problems, students were assessed using the unit test approximately every 2-3 weeks, depending on the length of the unit. This teacher also reported pairing higher performing students with struggling students in class during group work, as well as using leveled assignments to differentiate.

Treatment lessons were somewhat uniform between the second and fourth grade classrooms. One second and one fourth grade teacher reported using the Daily Spiral Review and Problem of the Day for review and to introduce the lesson, with the fourth grade teacher noting that these were used only sometime due to lack of time. In addition, the Problem of the Day was reported as not being used by the other second grade teacher because of lack of time. During math time, lessons would begin with Interactive Learning, explicitly stating the Purpose of the lesson as well as modeling multiple problems for the classes. Vocabulary was reviewed when applicable, and the Visual Learning Band was then discussed as a whole group. These preliminary activities usually took up the first 30-35 minutes of class, after which the teachers would use the Guided Practice before assigning students problems as part of the Independent Practice. The Quick Check was used as a short assessment for student understanding, with the exception of the fourth grade teacher (due to lack of time). The treatment classrooms attempted to use student center activities, though it was noted that there was not enough time in the schedule to do it regularly. The fourth grade treatment teacher would also differentiate assignments between her two leveled classes, incorporating the Practice Master worksheet with one group while using the Enrichment Master with the other.

The second grade teachers reported assigning homework twice a week in both classrooms, designed to take between 10-15 minutes, though one teacher explained that these assignments were given out on Monday and due at the end of the week. In contrast, homework was assigned to fourth grade treatment students almost every weeknight except for Fridays, and was checked

at the beginning of each math period. Both 2nd treatment teachers used Quick Checks, homework, and topic tests for assessing students. These two teachers met regularly to collaborate on the implementation of the enVisionMATH program.

Highlights – Overall, treatment teachers had a good grasp on the core elements of implementation. All teachers are quite experienced and know how to engage their students in order to make mathematics learning interesting. Across all classrooms, they engaged in problem-based activities and made modifications as needed in order to meet the needs of their students. There were also no notable differences between treatment and control students in terms of mathematics ability. There was also no evidence of contamination.

School G – Ohio

About the School: School G is an old public school located in a rural area of suburban Ohio. A modular trailer outside of the main building houses the two fourth grade classrooms. Due to the age of the building (listed on the National Register of Historic Buildings) and overcrowding issues, a brand new school is being built directly behind the current site. The school is situated in a lower middle class neighborhood, with 37% of students described as being economically disadvantaged, which is slightly higher than the state average. During the 2006-2007 school year the Ohio Department of Education classified 11% of the students as disabled and 7% as gifted. Enrollment for grades K-6 was 270 students. The student population is very limited in term of diversity as follows:

- White, non-Hispanic: 96%
- Unspecified: 4%

During the 2006-2007 school year, 100% of teachers were fully certified with at least a Bachelor's degree, and 63% had a Master's degree or higher. On average, teachers had 17 years of teaching experience behind them. There were approximately 17 students per full-time employee at this school.

Ohio uses the Performance Index as an overall indication of how well students perform on standardized tests each year. Scores are based upon how each student performs on all tested subjects in grades 3 through 8 and 10. The Performance Index ranges between 0 and 120, with 100 being the statewide goal for all students. In 2007, students at this school had an overall Performance Index score of 98. During that same year, 93% of second grade students performed at or above proficiency in math compared to the state average of 85%, and 94% of fourth grades students performed at or above proficiency in math compared to the state average of 76%.

Study Participants: There were 68 students in 4 classes at the school participating in the study, including 28 second graders and 40 fourth graders. There were two second grade teachers and two fourth grade teachers, one each for treatment and control. The average class size was 17.75 students, ranging from 15 students to 21 students per class, with larger class sizes in the fourth grades.

All teachers described their classes as being highly typical and comparable to other classes in the same grade at this school. Classes had a broad range of abilities, and teachers described their classes as being largely comprised of average students, with only a few low-performers and high-performers. All teachers also had at least one special education student, though one fourth grade treatment teacher, had a larger number of special need students. All teachers also had an assistant or inclusion tutor present during math lessons, primarily to help these struggling special education students. No significant behavior issues were observed and students generally appeared to be engaged in the lessons.

Math Curriculum and Resources: There were no school-wide initiatives or special programs that targeted math instruction. Control teachers used a variety of instructional materials from publishers including those that they selected, adapted, or developed themselves. The fourth grade control teacher included a commercial traditional basal math program (published 2001), while the second grade control teacher used a variety of materials, including a more investigative commercial text with hands-on activities correlated to state strands, teacher-produced student folders, various manipulatives and select pages of an additional commercial, traditional basal textbook.

While both treatment teachers primarily used the enVisionMATH program for their math lessons, the fourth grade treatment teacher also used some supplementary materials from lessons she had used prior to beginning enVisionMATH, such as a problem of the day by another resource, which she initially used but did not continue throughout the year, and was used in place of the enVisionMATH Problem of the Day. In addition, she also utilized the computer program *Blast-Off* for students to explore various math concepts independently.

Treatment teachers generally paced their classes as the lessons were organized in the enVisionMATH program, with some variation as state standardized testing approached in the spring to make certain that necessary topics were covered. Control teachers reported using the Ohio state standards and the correlating school progress reports to target math skills and instruction.

Resources for most classrooms were above average, and had anywhere from 3-8 computers available for student use. Computers only appeared to be present for teacher use in the treatment classrooms. Each room was also equipped with a television and VCR and some teachers had access to a SMART Notebook and ELMO.

Instructional Practices and Strategies: Math lessons were observed during all times of day, from 9:20 a.m. upon student arrival until 2:00 p.m. Second grade lessons were generally scheduled for 45 minutes per day, five days per week. Fourth grade lessons lasted slightly longer, and were 60-75 minutes per day, five days per week.

The fourth grade treatment class generally started her lesson with the Daily Spiral Review, which took approximately 10 minutes. During this time the teacher would take the opportunity to emphasize math vocabulary as she went over answers. This was followed by the Quick Check from the previous day's lesson, prior to going over the previous day's homework, which took about 20 minutes. During this time they often focused on problem-solving and worked in

partners on problems that everyone seemed to get stuck on. Next the students took out their notebooks and did the Problem of the Day, followed by cross checking answers with a partner. This took about 10 minutes. While going over these exercises, the teacher generally modeled problems, discussed answers with students and provided some real-life examples. The teacher then clearly stated the objective for the day's lesson and proceeded with the Visual Learning Band. Next the teacher would go through Guided Practice for approximately 15 minutes, followed by the start of Independent Practice. Overall her lessons tended to follow a format that was heavy on review.

In terms of her pedagogical approach the fourth grade treatment teacher was slightly more traditional, but incorporated some aspects of inquiry-based instruction. One exception is on Friday, when more activities and games take place. Because of pacing issues in the beginning, the fourth grade treatment teacher excluded components of the program that she did not feel were imperative and therefore her lessons adhered only somewhat to the enVisionMATH guidelines. During the Spring visit it was observed that she had worked out pacing issues and reported that she was using almost all of the essential components of the program with the exception of the Center Activities and Leveled Homework.

The fourth grade control class generally consisted of a lecture with discussion or some type of class problem-solving activity. The first five minutes of most lessons were devoted to working on problem-solving strategies. Next, vocabulary words were introduced and students would look up the definitions up in their textbook glossary and write down examples of when they might use the math vocabulary. The remainder of the class was spent working first as a class and then independently to finish the days work. For example, students might take out their textbooks and work together as a class on problem-solving and analyzing word problems followed by students working together in groups to finish the textbook questions or problem-solving puzzles. In general the format of the class involved good student-teacher and student-student collaboration and plenty of discussion. The overall pedagogical approach was centered in between inquiry-based and traditional.

The second grade treatment teacher generally began each lesson by problem modeling using e-tools with the SMART Notebook. The Interactive Learning Activity and Visual Learning Band was then introduced. This was followed by Guided Practice in the Student Workbooks completed as a class. Next, students moved on to the Independent Practice section while the teacher moved around the classroom checking student work. This was followed by the Problem of the Day which took about 5 minutes. The teacher would then go over the problem with class making sure to introduce specific vocabulary. As closure the teacher would generally lecture and model a few more problems before assigning Interactive Homework. Leveled Homework was not used and all students were assigned the same work. Students also kept their own math journal, in which Interactive Learning exercises were done. Overall her pedagogical approach was more traditional and textbook-based and she reported that she did not think students needed manipulatives in order to understand how to solve problems. The second grade treatment teacher reported that she did not use Error Intervention or Quick Check, but overall her lessons mostly adhered to the enVisionMATH guidelines.

The second grade control teacher generally began her lessons by doing a group problem which she problem modeled using the ELMO through the television, which took about five minutes. After this students were allowed to work with partners using homemade flashcards for about five minutes. Each student had several tools available for their use during the lesson, such as a 10s square, a 100s chart, a number line and manipulatives. She also integrated a few minutes at the beginning of the lesson to work on basic skills like skip counting by twos, fives, and threes using these various tools. Students then worked independently on worksheets for approximately eight minutes and then came together as a class to go over the work, which took the remainder of the class period. For example, each student was assigned one problem from the worksheet to work through in front of the class. Students were required to come up and do their problem on the board, explaining their steps and show their work. They were expected to use math vocabulary correctly to explain their steps. The teacher assigned the challenge problems from the worksheet to specific students, who then came up and worked through it. The unfinished work was assigned as homework. This teacher had many years of teaching experience and this was evident in the quality of lessons and instructional practices. Her pedagogical approach was slightly more inquiry-based with fundamentals of traditional teaching.

Highlights: Overall control and treatment teachers were equally capable in their teaching styles. As well, one fourth grade control teacher placed a great deal of emphasis on solving story problems and math reading. In terms of pedagogical approaches, teachers across both control and treatment were centered somewhere between inquiry based and traditional. All the teachers assigned homework on a regular basis. Student capabilities between treatment and control classes were generally similar and no behavioral problems were observed at all. There were no signs of obvious contamination between treatment and control teachers. Implementation levels for treatment teachers were moderate overall.

School H – Montana

About the School: School H is located in a small, rural town in western Montana. The school building houses grades K-6 and the building itself is older but well-kept and clean. On average there are two classes per grade level at the school and total enrollment is approximately 240 students. Because the town that the school serves is small, there is a great deal of community involvement in the school, and teachers and staff are very familiar with children's home and family situations. There is a great deal of evidence of student reward/incentive programs at the school and it seems clear that teachers and staff are focused on student success. The vast majority of students enrolled at the school are white. The ethnic breakdown of the school is as follows:

- White, not Hispanic: 97%
- American Indian/Alaskan Native 3%
- Black, not Hispanic: <1%

Additionally, more than half of the school's student population (57%) was eligible for free/reduced lunch. Class sizes at the school were relatively small and there were 13 students per

full-time teacher. Montana uses the Montana Criterion-Referenced Test (MCRT) to assess students in grades 3 though 8 and 10 in reading and math during the 2006-2007 school year. The MCRT is a standards-based test, which measures specific skills defined for each grade by the state of Montana. School H scored below the state average in both reading and math on the MCRT for 2006-2007. The third grade students in the school scored 66% proficient in math compared to the state average of 68% and fourth grade students scored 56% proficient compared to the state average of 67%.

***Study Participants*:** There were four classes and approximately 63 students participating in the first year of the enVisionMATH study; two second grade classes (n=31, one treatment and one control) and two fourth grade classes (n=32, one treatment and one control). The average class size was 16 students ranging from 15 to 18 students per class.

In general, teachers described their classes as average. However, both the second and fourth grade treatment teachers noted that their classes performed at a lower level than the control classes. This was particularly true of the second grade treatment class, where there were more low-performing students and students with behavioral problems than the control class. The treatment teacher noted that these class characteristics had an effect on her pacing and ability to cover material and move on quickly. It was much the same for the fourth grade classes where the treatment teacher noted that student motivation was an obstacle, whereas the control teacher pointed out that her students were easy to keep on task and that for the most part they covered material quickly. The fourth grade treatment teacher had a slightly different group of students for his math class than for homeroom instruction. The homeroom class included a combination third and fourth grade class, whereas, some of those third grade students rotated out for math instruction. The actual teaching ability of the participating study teachers in the school was comparable across all teachers. It was clear that all of the teachers genuinely cared about their students and wanted to give them the best math instruction possible. However, as noted, the differences in their students had an effect on the way they taught to their individual classes.

Math Curriculum and Resources: There were two control programs used during the first year of the enVisionMATH study at this school. In the second grade, the control class used a very traditional program with emphasis on review, daily drill and math practice. The program incorporated a spiraling approach to math instruction building one lesson at a time, and not in a traditional chapter format or "topic" based format like the enVisionMATH. While the program did offer some hands-on investigative learning opportunities, the focus was much more on teaching concepts and then learning them through constant review and practice. The control teacher used this program as her main resource. She strictly followed the teachers guide and completed lessons in the sequence of the program.

The control program used in the fourth grade was a text based program that also incorporated learning through hands-on activities. The structure of the program was similar in that a Problem of the Day and Spiral Review were a part of the program. The primary difference was that the enVisionMATH program had more emphasis on investigations based learning and visual examples to help reinforce topics. However, because this was an older program and components were not always available, the control teacher did a great deal of supplementing with her own materials and activities rather than using her assigned program exclusively. She used the

program as her sequence guide of what to cover, but did not follow lessons from the text verbatim. She supplemented with her own teacher created activities, and also some worksheets and practice sets from other publishers.

Treatment teachers at the school used the enVisionMATH program as their primary resource for math instruction with very little supplementing. Teachers used the core components of the enVisionMATH program as outlined in the implementation guidelines. Both treatment teachers did note supplementing the program from time to time with facts practice worksheets or flashcards, but this was typically done at a time separate from their enVisionMATH lesson time. For example, the second grade treatment teacher also had a separate morning math time that included calendar work, a facts practice review and some board work. The second grade teacher also used a technology program with practice and games for early finishers for independent practice. Students would go back to the computers from 5-10 minutes and work while other students finished the day's math lesson assignment.

***Instructional Practices and Strategies*:** Math lessons were taught at varying times throughout the day. The second grade control teacher taught her math lesson earlier in the day, while the treatment teacher taught math in the afternoon following lunch. At the fourth grade, it was the opposite, with the treatment class earlier in the day and the control class in the afternoon just before release for the day. As noted, the fourth grade treatment teacher had a homeroom class that was a third and fourth grade combination class. The third grade students rotated out for math. Because of this rotation schedule, math time was very consistent through out the year. The average length of math lessons was 55 minutes, ranging from 50-60 minutes per lesson.

The typical class structure for math lessons varied across the participating classrooms. As noted previously, the second grade control teacher followed her assigned program strictly, while the fourth grade control teacher did not. However, the basics of math lesson format followed similar patterns. For the second grade class, lessons began with a 5-10 minutes warm-up completing a fact practice sheet. This was followed by teachers' lesson instruction, including review of anything from the previous lesson. The lesson concluded with independent practice worksheets which were completed in class with the remainder assigned as homework. A typical lesson for the fourth grade control teacher began with 10-15 minutes of warm-up activity where students were in teams to complete worksheets. This was then followed by the core control program lesson instruction and teaching new math, which was followed by independent practice.

Both of the treatment teachers used the enVisionMATH program as their main math instruction resource. However, the second grade treatment teacher tended to do some reordering of the main program components to help her better meet the needs of her class. For example, sometimes she completed the Visual Learning Band as a post-instruction review rather than when it appeared in the lesson. She also did the Problem of the Day and Spiral Review during her morning math time rather than as the warm-up for her actual enVisionMATH lesson. The fourth grade teacher rarely used the Problem of the Day or Spiral Review (note that the trainer suggested this in order to assist with pacing issues that this teacher was experiencing). Instead of the those two activities, the teacher began classes with the Quick Check review as a warm up, which took about 10 minutes. Neither of the teachers regularly used the Diagnosis and Intervention System.

Both teachers attributed this to lack of time and not having a great understanding of how the system worked.

In terms of homework, the fourth grade treatment teacher indicated that he rarely assigned homework in math. During an interview, it was noted that this was due to a new initiative at the school for no homework to be assigned to students. However, all of the other participating teachers, both treatment and control, noted that they assigned some homework in math. The control teacher said she assigned homework a few nights a week, but it was never more than 5-10 minutes per night. In the second grade treatment class, the teacher noted that she assigned homework typically every night except over the weekend. The second grade control teacher consistently assigned 5-10 minutes of homework per night; students were assigned to complete a worksheet from class.

For assessment, all teachers reported using constant informal assessment to check on student's progress and ensure that students were keeping up. For formal assessments, teachers used the assessments provided in their assigned programs.

Aside from program-based differences, the comparability of all classes was somewhat similar. All teachers taught new math content, and then followed-up with review and practice of those concepts. Treatment classes tended to have somewhat more investigative based activities, with students explaining how they arrived at a solution. However, this was also evident in the fourth grade control classroom, just to a lesser extent.

Highlights: All of the teachers at the school were equally comparable in their ability to teach math. There were no major behavioral issues observed and teachers did a good job of working with and engaging students in math instruction. The most notable difference between the treatment and control classrooms was that the program used in the second grade control class distinctly contrasted the enVisionMATH program. Initially, the treatment teachers experienced some pacing troubles as they and their students adjusted to the new enVisionMATH program. The treatment teachers indicated that their pacing issues were resolved once they achieved a certain comfort level with teaching the program and their students got used to it.

Appendix C:

Key Features and Resources for Treatment and Control Programs

Key Program Features	enVision MATH (2009)	Control Program #1 (2004): Site A (2nd and 4th grade)	Control Program #2 (1998): Site B (2nd and 4th grade)	Control Program #3 (2005): Site C (2nd and 4th grade)	Control Program #4 (2005) Site D and E (2nd and 4th grade)	Control Program #5 (2007): Site F (2nd and 4th grade)	Control Program #6 (NA): Site G (2nd and 4th grade)	Control Program #7 (2000): Site H (2nd grade only)	Control Program #8 (2000): Site H (4th grade only)
▪ Focus is placed on daily, problem- based Interactive Learning that is reinforced through Visual development of concepts to gain deeper understandings and increased visual/verbal connections to math skills. ▪ Structured lessons provide step-by-step instruction that is easy to follow and helps build students' understanding ▪ Incorporates numerous hands-on activities for students ▪ Provides leveled practice and stimulating activities in every lesson to customize instruction to match students' abilities and interests. ▪ Developed so all of the lessons in the program can be taught before the time of the state test ▪ Each lesson begins with a Daily Review to allow teachers to gauge student understanding ▪ Problem-solving instruction is systematic and explicit ▪ Diagnosis and Intervention System allow for informed intervention and differentiation ▪ Provides ample opportunities for test preparation and review of previously learned concepts/skills		▪ Focus is on activity based mathematics ▪ Encourages students to think creatively, develop and articulate their own problem-solving strategies ▪ Includes ongoing and periodic assessment ▪ Organized into Units intended to link together for complete math instruction	▪ Focus is placed on problem-solving and increasing students critical thinking and understanding of math ▪ Each lesson begins with a clear "Objective" ▪ Includes visual models in each lesson to support concepts ▪ Every lesson includes informal assessment, and new vocabulary ▪ Incorporates suggestions for differentiating instruction	▪ Focus is placed on development of problem-solving skills ▪ Included embedded assessment and differentiated exercises and helps for teachers to gauge student understanding ▪ Incorporates hands-on activities for students Problem-solving instruction is systematic and explicit. ▪ Reading and writing connections and opportunities are provided throughout.	▪ Program designed to cover state and national standards giving teacher flexibility customize for state/local standards ▪ Focus is placed on building skills through conceptual understanding ▪ Emphasis on problem-solving and logical reasoning ▪ Built in assessment	▪ Focus on developing skills over time in a variety of contexts ▪ Provides real-world and cross-curricular applications for math ▪ 3 part lesson structure places emphasis on practice and differentiation options ▪ Includes ongoing, periodic formative and summative assessment options	▪ Various publisher programs ▪ Miscellaneous math program worksheets and student activity packets ▪ Teacher collected supplemental resources	▪ Focus on an incremental, spiraling approach to math instruction ▪ Review and practice is incorporated in each lesson ▪ Emphasis is on building students confidence in math ▪ Incorporates hands-on learning with a lot of independent practice	▪ Focus on learning math through skill practice and problem-solving ▪ Each lesson includes suggestions for differentiation and checking understanding ▪ Engaging warm-up or hands-on activity begins each lesson ▪ Embedded assessment included with the program

		enVisionMATH (2009)	Control Program #1 (2004): Site A (2nd and 4th grade)	Control Program #2 (1998): Site B (2nd and 4th grade)	Control Program #3 (2005): Site C (2nd and 4th grade)	Control Program #4 (2005) Site D and E (2nd and 4th grade)	Control Program #5 (2007): Site F (2nd and 4th grade)	Control Program #6 (NA): Site G (2nd and 4th grade)	Control Program #7 (2000): Site H (2nd grade only)	Control Program #8 (1998): Site H (4th grade only)
Content*										
2nd grade	Understanding Addition and Subtraction	Yes	Yes	Yes	NC	Yes	Yes	NC	NA	
	Addition Strategies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	
	Subtraction Strategies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	
	Place Value: Numbers to 100	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	
	Counting Money	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	
	Mental Addition	Yes	NC	NC	Yes	NC	Yes	NC	NA	
	Mental Subtraction	NC	NC	NC	Yes	NC	Yes	NC	NA	
	Adding Two-Digit Numbers	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	
	Subtracting Two-Digit Numbers	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	
	Using Addition and Subtraction	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	
	Geometry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	
	Fractions	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	
	Measurement: Length and Area	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	
	Measurement: Capacity and Weight	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	
	Time and Temperature	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	
	Graphs and Probability	Yes	NC	Yes	Yes	NC	Yes	Yes	NA	
	Numbers and Patterns to 1,000	Yes	NC	Yes	Yes	Yes	Yes	Yes	NA	
	Three-Digit Addition and Subtraction	Yes	Yes	Yes	Yes	NC	Yes	Yes	NA	
	Multiplication Concepts	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	
	Division Concepts and Facts	NC	Yes	Yes	NC	Yes	Yes	Yes	NA	

NC = No chapter devoted to this, touches upon throughout others
NA = This program was not used in this grade level
*Note that control teachers incorporated other program materials (other than the primary program). This comparison is based on content taught across all programs used.

		enVisionMATH (2009)	Control Program #1 (2004): Site A (2nd and 4th grade)	Control Program #2 (1998): Site B (2nd and 4th grade)	Control Program #3 (2005): Site C (2nd and 4th grade)	Control Program #4 (2005) Site D and E (2nd and 4th grade)	Control Program #5 (2007): Site F (2nd and 4th grade)	Control Program #6 (NA): Site G (2nd and 4th grade)	Control Program #7 (2000): Site H (2nd grade only)	Control Program #8 (1998): Site H (4th grade only)
Content*										
4th grade	Numeration	NC	NC	Yes	Yes	NC	Yes	NA	NC	
	Adding and Subtracting Whole Numbers	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	
	Multiplication Meanings and Facts	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	
	Division Meanings and Facts	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	
	Multiplying by 1-Digit Numbers	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	
	Patterns and Expressions	Yes	NC	Yes	NC	Yes	Yes	NA	NC	
	Multiplying by 2-Digit Numbers	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	
	Dividing by 1-Digit Numbers	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	
	Lines, Angles, and Shapes	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	
	Understanding Fractions	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	
	Adding and Subtracting Fractions	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	
	Understanding Decimals	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	
	Operations with Decimals	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	
	Area and Perimeter	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	
	Solids	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	
	Measurement, Time, and Temperature	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	
	Data and Graphs	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	
	Equations	Yes	NC	Yes	NC	Yes	Yes	NA	Yes	
	Transformations, Congruence, and Symmetry	Yes	Yes	NC	Yes	Yes	Yes	NA	NC	
	Probability	NC	NC	Yes	Yes	Yes	Yes	NA	Yes	

NC = No chapter devoted to this, touches upon throughout others
NA = This program was not used in this grade level
*Note that control teachers incorporated other program materials (other than the primary program). This comparison is based on content taught across all programs used.

Program Materials	enVisionMATH (2009)	Control Program #1 (2004): Site A (2nd and 4th grade)	Control Program #2 (1998): Site B (2nd and 4th grade)	Control Program #3 (2005): Site C (2nd and 4th grade)	Control Program #4 (2005) Site D and E (2nd and 4th grade)	Control Program #5 (2007): Site F (2nd and 4th grade)	Control Program #6 (NA): Site G (2nd and 4th grade)	Control Program #7 (2000): Site H (2nd grade only)	Control Program #8 (1998): Site H (4th grade only)
	▪ Topic Teacher Editions ▪ Teacher Resource Masters ▪ Overview and Implementation Guides ▪ Student Editions ▪ Interactive Homework Workbook ▪ Interactive Math Series Big Book (K – 2) ▪ Math Diagnosis and Intervention System ▪ Individual Student Manipulative Kits ▪ Teacher Overhead Manipulative Kits ▪ Center Activities Kits ▪ Visual Learning Bridge Transparencies ▪ MathStart Books by Stuart Murphy (K – 2) ▪ World Scapes Books (3 – 6) ▪ ExamView CD-ROM ▪ ETools CD-ROM	▪ Student Materials Kit ▪ Teacher Resources ▪ Math Curriculum Units	▪ Student Edition ▪ Teacher's Edition ▪ Literature/Trade books ▪ Supplemental worksheets ▪ Performance Assessments ▪ Language Resources ▪ Problem of the Day ▪ Daily Review Booklet ▪ Classroom/ Overhead Manipulatives ▪ Activity Kits ▪ Internet Resources	▪ Student Edition Textbook ▪ Teacher's Edition ▪ Teacher's Resource Package ▪ Homework Workbooks ▪ Transparencies ▪ Problem of the Day ▪ Math Vocabulary Kit ▪ Leveled Literature Library ▪ Manipulative Kits ▪ Technology package	▪ Student Textbook ▪ Student workbooks ▪ Teacher's Edition ▪ Teacher's resource book ▪ Problem-solving and intervention resources ▪ Manipulatives ▪ Electronic Resources (CD-ROM)	▪ Student Edition ▪ Teacher's Edition ▪ Student Math Journal ▪ Differentiation and Assessment Resources ▪ Technology components	▪ Various publisher programs ▪ Miscellaneous math program worksheets and student activity packets ▪ Teacher collected supplemental resources	▪ Teacher's Guide ▪ Teacher's Resource Package ▪ Manipulative kits ▪ Differentiated instruction guides	▪ Student Edition Textbook Teacher's Edition ▪ Teacher's Resource Package ▪ Transparencies ▪ Problem of the Day ▪ Math Vocabulary Kit ▪ Manipulative Kits

Appendix D:

Use of enVisionMATH Resources and Teacher Ratings of Usefulness

Table D1. Percent* of Teachers Using enVisionMATH Print Materials on a Given Month

Components	2nd Grade	4th Grade
Topic teacher Editions	69.2%	93.6%
Teacher Resource Masters	41.4%	65.6%
Overview and Implementation Guidelines	28.4%	48.9%
Student Editions	76.9%	95.5%
Reteaching Exercises in SE	41.0%	47.1%
Homework Workbook	62.4%	66.0%
Math Series Big Book	72.9%	0.0%
Math Diagnosis and Intervention System	16.9%	13.4%
Individual Student Manipulative Kits	64.8%	66.3%
Teacher Overhead Manipulative Kits	24.3%	26.9%
Center Activities Kits	18.6%	14.7%
Visual Learning Bridge Transparencies	56.3%	49.6%
Math Start by Stuart Murphy (2nd grade only)	24.2%	NA
WorldScapes Books (4th grade only)	NA	1.1%
Pouch – Center Activity Masters	37.2%	38.8%
Pouch – Home/School Connection Letter	21.9%	31.1%
Pouch – Vocabulary Cards	58.8%	26.7%
Pouch – Reteaching Worksheet	72.9%	76.4%
Pouch – Practice Worksheet	92.7%	77.9%
Pouch – Enrichment Worksheet	72.3%	60.1%

*% reflects percent of teachers who reported using the listed program materials anytime during the month as noted on the teacher log.

Table D2. Percent* of Teachers Using enVisionMATH Technology

Material	2nd Grade	4th Grade
Exam View Test Generator	1.0%	0.0%
E-TOOLS CD ROM	3.0%	24.1%

*% reflects percent of teachers who reported using the listed program materials anytime during the month as noted on the teacher log.

It should be noted that the low percentages of use for technology components may be due to the later arrival of these components.

Table D3. Percent and Rating Given to Usefulness of enVisionMATH Resources/Program Components

	Percent Who Rated Component Good to Excellent	Mean**	Std. Deviation
Topic Teacher Editions	84.2%	4.18	0.69
Teacher Resource Masters	88.9%	4.17	0.62
Overview and Implementation Guides	58.8%	3.53	0.62
Student Editions	79.0%	4.16	0.77
Interactive Homework Book	66.7%	3.80	0.68
Interactive Math Series Big Book (2nd grade only)	50.0%	3.50	0.93
Math Diagnosis and Intervention System	54.6%	3.73	0.79
Individual Student Manipulative Kits	63.2%	3.84	0.90
Teacher Overhead Manipulative Kits	75.0%	4.13	1.13
Center Activity Kits	50.0%	3.50	0.73
Visual Learning Bridge Transparencies	66.7%	3.88	0.86
MathStart Books by Stuart Murphy (2nd grade only)	100.0%	4.17	0.41
Problem of the Day	89.5%	4.16	0.77
Daily Spiral Review	89.5%	4.32	0.69
Interactive Learning (IL) activity	72.2%	3.81	0.79
Vocabulary words	77.8%	4.11	1.08
Guided Practice	95.0%	4.35	0.65
Independent practice (including Leveled homework)	94.7%	4.32	0.61
World Scape Books (4th grade only)	Not Rated by Teachers		
Exam View Test Generator CD ROM	Not Rated by Teachers		
eTools CD ROM	80.0%	4.00	0.79
Assessments that came with program	47.1%	3.47	0.72

*Only includes teachers who used the listed program components. Based on scale of 1-Poor to 5-Excellent.

Appendix E:

enVisionMATH
Implementation Guidelines

enVisionMATH Research Study Implementation Guidelines

Introduction

Welcome, and thank you for participating in the research study being conducted by PRES Associates on Scott Foresman-Addison Wesley *enVisionMATH*. We believe your experience with our study will be rewarding and enjoyable. Not only will you contribute to cutting-edge research, but you will also benefit from first-rate professional development provided by Pearson Scott Foresman professional training specialists.

We understand that it may be challenging to change former practices and implement a new math program. Therefore, we greatly appreciate the time and effort you will be putting into making this study a success. However, we also realize that there may be obstacles and challenges as you begin to implement this program. Under these circumstances, we want and need to hear from you; we will guide you through those challenges. In fact, it is critical that any problems you encounter be addressed as soon as possible to ensure that this program is being implemented to its full potential. Feel free to contact PRES Associates via e-mail at studies@presassociates.com if you have any questions, problems or concerns.

The following provides answers to some common questions teachers may have related to this study. Please read through all of these questions/answers. Again, should you have further questions, please contact PRES Associates.

Why Is This Research Being Done?

As you are aware, the No Child Left Behind Act (NCLB) of 2001 requires that educational materials and strategies used by educators in the classroom *must be proven by scientific research to improve student achievement in the classroom.* Pearson Scott Foresman has developed a strong research model for determining that their programs are scientifically-based. As part of this ambitious research agenda, Pearson Scott Foresman has contracted with PRES Associates[35], an external educational research firm, to conduct a rigorous quantitative research study on the effectiveness of the Scott Foresman-Addison Wesley *enVisionMATH* program. This study will contribute to the growing research base behind Pearson Scott Foresman math programs and the effectiveness of different approaches to math instruction.

Why Do I Need Professional Development?

It takes more than a good curricular program to raise students' knowledge of mathematics. It also takes good teachers with a thorough understanding of the curriculum, who are supported by professional development, school administrators, and parents/guardians. To this end, it is

[35] PRES Associates is an external, independent, educational research firm with an established track record in conducting large-scale, rigorous evaluations on the effectiveness of research materials.

hoped that through the professional development training sessions provided by Pearson Scott Foresman on the use of its elementary math program, all teachers participating in the study will gain the knowledge and skills to successfully implement this program right from the start. As you will soon learn, this math program provides numerous teaching resources and supports. In order to implement this program successfully, it is essential that teachers have a thorough understanding of the resources provided by the *enVisionMATH* program. Rather than having teachers figure it out on their own, professional trainers will guide you through this process, offering examples of when to use certain materials, how to manage and supplement classroom instruction, what types of assessments to administer, and so forth.

Why Do I Need To Follow These Implementation Guidelines?

The Teacher Implementation Guidelines were developed by the senior author of the *enVisionMATH* program and clearly outline the essential components of the program that must be implemented in order for the program to be maximally effective. The guidelines are designed for teachers to use when implementing the new program in their "treatment" class(es). The guidelines point out key program components that *must* be implemented during math lessons. These key program components have the greatest influence on student learning and performance, and therefore should be implemented. In addition, it is critical to ensure that all teachers are implementing a similar instructional model. That is, if teachers are modifying the program to an extent that it no longer resembles the original program, the study is no longer an accurate evaluation of the *enVisionMATH* program. In sum, by providing these implementation guidelines, we are attempting to (1) maximize the potential of this math program to help your students, and (2) ensure that the program is being implemented with fidelity across all teachers assigned to use this program. To reiterate, *it is essential that all teachers implement the program fully in their "treatment" classes as prescribed in the following implementation guidelines.* That being said, there are optional parts to the program as well as ancillary materials that provide you with the flexibility you need to address unique student needs or contexts. *We trust your professional judgment and ask that you try to implement the program as best you possibly can while meeting your instructional needs.*

Again, thank you for your participation in this study. You are an integral part of this study and we appreciate your assistance. We look forward to working with you.

Guidelines for Implementing Scott Foresman's *enVisionMATH*

Randy Charles
Senior Author

PACING

enVisionMATH is organized into 20 content topics. Each topic focuses on a particular strand that can be identified by color (e.g., blue is number and operations). Each topic contains 4 to 9 lessons and develops one or a few related content standards in depth.
Unlike other programs you have used in the past, *enVisionMATH* was developed so all of the lessons in the program can be taught before the time of the state test. In the past, teachers often jumped around in a textbook to make sure that all key topics were developed before the test; this jumping around is no longer needed with *enVisionMATH*. So, with this in mind, watch that you maintain a pace that will enable you to teach all 20 topics before the state test or before the end of the year if there is no test at your grade.

REORGANIZING THE PROGRAM

The Teacher's Edition for *enVisionMATH* is 20 separate booklets each addressing one content topic as described above. Also, there is an Overview and Implementation booklet for the program. This organization makes it easy for you top teach topics in any order you wish. However, the sequence of 20 topics in *enVisionMATH* is the default sequence; it is a mathematically sound sequence that makes all needed connections and develops content in a planned way across the year. You can start with Topic 1 and teach in order through Topic 20. Or, you may want to teach topics in a different order. For example, suppose measurement is a challenging topic for your students and you would like to teach that earlier in the year. You can change the order of the topics in *enVisionMATH* but you must be careful. As you know, many ideas in mathematics must be developed in a sequential manner. For example, place value to the millions place should be developed after place value to the thousands place. If you want to teach topics in a sequence different from the default sequence, we have provided guidelines for doing this (see the topic clusters documents). If you do make changes in the sequence using the guidelines provided, remember that all 20 topics should still be taught.

PREPARING TO TEACH THE TOPIC and LESSON

1) Be sure to review all of the material in the front matter of each topic, that is, the pages that precede the lessons. Pay particular attention to the Topic Planner as that shows how lessons connect across the topic. The Math Background for Teachers may also be helpful to gain an understanding of how content is developed in the program.
2) Familiarize yourself with the lesson using the **Quick and Easy Lesson Overview.** The **Objective** describes what students should be able to do at the end of the lesson. The **Essential Understanding** describes what students should understand at the end of the lesson. Note that while each lesson has a unique objective, the essential understanding is often developed in more than one lesson.
3) Use the **Mathematics Background** as a quick review for you of the mathematics content underlying this lesson. Go back to the Mathematics Background section in the Topic overview as needed for additional support.
4) There are four parts to the instructional plan used in *enVisionMATH*. Each of these is described below.
    (1) Daily Review
    (2) Develop the Concept: Interactive

(3) Develop the Concept: Visual
(4) Close/ Assess and Differentiate

TEACHING A LESSON: A Four-Part Instructional Model

Part 1. *Daily Review* and *Problem of the Day*. (see the Teacher's Edition)
The *Daily Review* and the *Problem of the Day* review key skills and concepts. You must use the Daily Review for this study; the Problem of the Day is optional. Note that these can be used at any time in the day.

Part 2. Develop the Concept : Interactive

a. Instruction on the new material of the day should begin with the *Interactive Learning* (IL) activity in the teacher's edition (TE). This should be done before opening the pupil book at grades 3 and up. In the primary grades, a recording page supporting the Interactive Learning activity is the first page of the pupil book sheet. At the primary grades, the IL might last approximately 15-20 minutes; in the upper grades the IL might last 10-15 minutes.
b. The focus of the Interactive Learning activity should be on developing students' understanding of mathematics. Start by using the **Engage** with the entire class to make connections. **Set the Purpose** connects the new concept to be learned to previously learned concepts. **Connect** relates the new concept to be learned to something in the real-world.
c. Interactive Learning activities are problem-based activities. There is a four-part teaching approach that should be used for these.
    1) Start by Posing a problem
        a) The problem will require some thinking.
        b) Students will have to grapple with ideas.
        c) Many problems will be able to be solved in more than one way.
    2) Allow time for students to solve the problem.
        a) Usually have students work in pairs or groups.
        b) Manipulatives may or may not be called for.
        c) You should walk around the room and observe and facilitate as students work.
    3) Have students share their thinking and work.
        a) Have alternative solutions or ways of thinking shared.
        b) Have students explain their thinking and work.
    4) End by making the important mathematics explicit by connecting to the students' thinking and work. (Teacher-directed instruction). See the teaching notes in the Interactive Learning for ways to do this.

Each IL starts by posing a problem for students to solve. Use the questions in the Teacher's Edition to facilitate student's work. After working on the problem have students share their thinking and work. Then your role is to make the important mathematics of the lesson explicit usually by modeling for students. New vocabulary should also be introduced at this time. Make sure the students recognize what mathematical understanding they were supposed to take away from the IL. Teaching strategies used here are called out as bold phrases in the margin (e.g., **Model, Academic Vocabulary, Use Drawings**).

Part 3. Develop the Concept: Visual

a. Start by telling students what they will be learning that day; see the **Set the Purpose** statement in the Teacher's Edition (TE). Usually this is just a reminder of what they were told when setting the purpose for the Interactive Learning activity.

b. Every pupil edition lesson has a Visual Learning Bridge (VLB) across the spread of pages. The VLB connects the work the students did in the Interactive Learning to a visual presentation in the pupil's edition of the math concepts and skills. Many lessons in the Pupil Edition are "hands on" lessons as indicated by the logo at the top of the Pupil page. Use manipulatives if they are called for.

c. There are 3 options for using the VLB.

(1) Use the VLB in the PE only. Have students look at the PE page and use the TE questions to walk through the VLB.

(2a) Use the VLB transparency with the pupil edition (PE) open. This is really the same as (1) but the transparency allows you to direct attention to the correct part of the PE, and it will make it easier to ask questions since they are on the transparency.

(2b) Use the VLB transparency with the PE closed. This allows you to have a teacher-directed conversation but it can be varied. Since can show one frame at a time, the transparency can sometimes be used for another problem-based learning task. The problem can be presented and questions asked. Then students can work together to solve the problem. They can share their work. Then the teacher can walk through the remaining parts of the transparency.

(3a) Using the **Visual Learning Animation (VLA)** with the sound off. This is the same as (2b) as you can stop and ask questions when you like and it can also be used with a problem-based approach as above,

(3b) Using the VLA with the sound on. This can be used in a variety of ways. It can be a problem-based approach by stopping and having the students solve the problem or the teacher can walk through the example asking questions.

d. Use the **Guided Practice** to assess students before assigning **Independent Practice.** Use the **Error Intervention** ideas and follow up **Reteaching** as needed.

e. Assign appropriate **Independent Practice** exercises, and monitor students as they work on these.

Part 4. Close/Assess and Differentiate

a. Bring closure to the lesson using the **Close** in the TE. Use this statement to make explicit to students what they should have understood from today's lesson.

b. Use the **Quick Check** to assess students' understanding. Use the scoring rubric provided to make decisions about what if any follow-up instruction or practice is needed.

a. An **Intervention** teacher-directed learning experience is provided for those students needing additional instruction under the guidance of the teacher.

b. While the teacher is working with those students needing additional instruction, other students can further explore the ideas in the lesson using the On-Level and Advanced **Center Activities.**

c. If needed, assign the **Leveled Homework** as appropriate.

PRES Associates

PO Box 10730
Jackson, WY 83002
Phone: 307-733-3255, Fax: 307-733-2017
info@presassociates.com