Educational Research Institute of America
www.eriaonline.com • email: info@eriaonline.com
An affiliate of Beck Evaluation & Testing Associates, Inc.

A LONGITUDINAL ANALYSIS OF STATE MATHEMATICS SCORES FOR FLORIDA SCHOOLS USING *SAXON MATH*

Report Number 365

May 2009

Contents

A LONGITUDINAL ANALYSIS OF STATE MATHEMATICS SCORES FOR FLORIDA SCHOOLS USING *SAXON MATH*

This report describes a three year longitudinal study of the instructional effectiveness of *SAXON MATH*, a mathematics program designed for use in kindergarten through grade 12.

Project Background

We live in a mathematical world. Never before has the workplace demanded such complex levels of mathematical thinking and problem solving (National Council of Teachers of Mathematics, 2009). Clearly, those who understand and can do mathematics will have increased opportunities in the workplace. Mathematical competence can open doors that will allow for educational and career advancements. A lack of mathematical competence can close those doors.

Unfortunately, in terms of mathematical skills, the United States is quickly falling behind the rest of the developed world. A recent study comparing the math skills of students in industrialized nations found that U.S. students in grades 4 and 8 consistently performed below most of their peers around the world, a trend that continues into high school (Mullis, Martin, Gonzalez, & Chrostowski, 2005). And although the latest results from the National Assessment of Educational Progress (2007) showed improvements in the math performance of students in grades 4 and 8 nationally, upon closer examination, only fourteen of the fifty states showed improved scores at both grade levels. Seventeen states did not show improvements at either grade level. Further, low-income and minority students in the U.S. perform relatively poorly in math as early as kindergarten and first grade (Denton & West, 2002). By the third grade, the number of American students showing signs of math learning difficulties increases significantly (Ostad, 1998, 1997; Geary, Hoard, Byrd-Craven, & DeSoto, 2005).

To address concerns that many students lack essential skills to be successful in mathematics-related careers, President Bush called for the creation of a National Mathematics Advisory Panel in April, 2006. This panel was charged with fostering "greater knowledge of and improved performance in mathematics among American students" (U.S. Department of Education, 2008).

On March 13, 2008, the National Mathematics Advisory Panel submitted its final report. In the report, the Panel stressed how critical it is that students succeed in algebra, in part because doing so will make them much more likely to succeed in college and be prepared for better career opportunities in the global economy of the 21st century. The Panel also emphasized the importance of children having a strong base in mathematics. Research shows that a strong start can be a major contributor to preventing later difficulties in math learning. Efforts must begin in early childhood, with a particular focus on the foundational skills learned from kindergarten through third grade. Effective early math education can help students to:

- Acquire the foundational knowledge and skills that they will need to be successful with algebra and other advanced math courses (National Association for the Education of Young Children and National Council of Teachers of Mathematics, 2002);
- Avoid retention in the early years by increasing math skills (Magnuson, Myers, Ruhm, & Waldfogel, 2003); and
- Develop positive attitudes toward learning math early on (Ma, 2000).

There has never been a greater need to ensure that the math programs today's young students are using are optimally supporting them in developing the mathematical skills and strategies required for success in high school, in college, and in the workplace. Because of the importance of determining the effectiveness of programs designed to support young children with mathematics instruction, Houghton Mifflin Harcourt contracted with the Educational Research Institute of America (ERIA) to study the effectiveness of the *SAXON MATH* program. This report presents the findings from that study.

Research Question

The following research question guided the design of the study:

> Is *SAXON MATH* instructionally effective in improving students' mathematical skills and strategy use over time?

Design of the Study

A quasi-experimental, pretest/posttest design was used for this study. Florida schools using the *SAXON MATH* program at grades 3 and/or 5 during the 2005-2006, 2006-2007, and 2007-2008 school years were included in the study. Achievement data from the spring 2005 and the spring 2008 administrations of the math portion of the Florida Comprehensive Assessment Test (FCAT) were used as the pretest and posttest respectively.

In order to identify Florida *SAXON MATH* schools for inclusion in the study, Houghton Mifflin Harcourt provided researchers with a list of Florida elementary schools that had purchased *SAXON MATH*. Researchers then telephoned the administrators at each of these schools to determine the year each one had started using *SAXON MATH* at grades 3 and 5 and for how long each had continued to use the program at those same grade levels. This information was not readily available from all school administrators due to changes in administration and/or lack of records. Also, some schools were unwilling to provide the requested data. Schools were included in the study if it could be verified, based on these phone calls, that they had started using the program at grade 3 and/or grade 5 no later than the 2005-2006 academic year and had continued to do so through the 2007-2008 academic year or longer.

A total of 22 schools in Florida were verified as having used *SAXON MATH* at grade 3 from the 2005-2006 academic year through the 2007-2008 academic year. A total of 14 Florida schools were verified as having used *SAXON MATH* at grade 5 for the same three-year period. Only nine of the schools are the same across the grade 3 and grade 5 lists of schools. The differences are due to the fact that some schools did not adopt the program at all grade levels during the same academic year but instead adopted the program at one or two grade levels each year over several years until the program was implemented at all grade levels.

For each verified school, researchers downloaded the FCAT mathematics data that was available to the public from the Florida State Department of Education (FSDOE) Web site. For each administration of the FCAT, the FSDOE Web site provides the percentage of students at each school achieving at each of five performance levels on the math portion of the FCAT, as well as the mean standard scores on the math portion of the FCAT for each school.

Instructional Approach under Study

The description of *SAXON MATH* provided by the publisher states the following:

> A well-articulated curriculum challenges students to learn increasingly more sophisticated mathematical ideas as they continue their studies. John Saxon had a similar philosophy in mind when in the early 1980s he developed his theory-based distributed approach to mathematics instruction, practice, and assessment. Utilizing this approach, the *SAXON MATH* K–12 program was created with a comprehensive approach to mathematics. Because smaller pieces of information are easier to teach and easier to learn, the *SAXON MATH* series was developed by breaking down complex concepts into related increments. The instruction, practice, and assessment of those increments were systematically distributed across each grade level. Practice is continual, and assessment is cumulative.
>
> The *SAXON MATH* approach differs from most programs in that it distributes instruction, practice, and assessment instead of massing these elements throughout the lessons and school year. In a massed approach, instruction, practice, and assessment of a skill or concept occur within a short period of time and are clustered within a single chapter or unit. In the *SAXON MATH* program, as students encounter new increments of instruction, they are also continually reviewing previously introduced math concepts. Frequent assessments of newer and older concepts are encountered throughout the lessons, ensuring that students truly integrate and retain critical math skills.

Description of the Research Sample

A total of 22 schools in Florida were verified as having used *SAXON MATH* at grade 3 from the 2005-2006 academic year through the 2007-2008 academic year. Fourteen schools were verified as having used the program at grade 5 for the same period. Tables 1 and 2 provide demographic summaries of the schools included in the study at each grade level. The average enrollment for the grade 3 schools was 563. The average percent of students enrolled in free and reduced lunch programs across the grade 3 schools was 49%. The average percent of minority students was 31%. For the grade 5 schools, the average enrollment was 559 students with an average of 43% of the students enrolled in free/reduced lunch programs and an average of 30% of the students identified as minority across the schools.

Table 1
Demographic Characteristics of Florida *SAXON MATH* Grade 3 Schools Included in the Study

School	***Grades***	***Locale***	***Enrollment***	***% Free/Reduced Lunch***	***% Minority***	***% Limited English Proficient***
1	K to 5	Urban	371	53%	34%	11%
2	K to 5	Suburban	842	43%	10%	0%
3	K to 5	Suburban	606	56%	22%	1%
4	K to 5	Urban	290	85%	80%	0%
5	K to 6	Suburban	625	59%	43%	1%
6	K to 5	Urban	91	45%	30%	1%
7	K to 5	Suburban	488	35%	26%	1%
8	K to 5	Suburban	485	45%	10%	1%
9	K to 5	Urban	897	34%	14%	1%
10	K to 5	Suburban	508	46%	38%	1%
11	K to 5	Suburban	99	19%	36%	4%
12	K to 5	Urban	672	40%	20%	1%
13	K to 5	Suburban	653	84%	30%	10%
14	K to 5	Suburban	883	34%	9%	1%
15	K to 5	Urban	436	89%	70%	1%
16	K to 5	Rural	698	54%	50%	5%
17	K to 5	Urban	800	25%	22%	1%
18	K to 5	Rural	734	38%	23%	1%
19	K to 5	Suburban	414	41%	9%	1%
20	K to 5	Urban	379	56%	52%	2%
21	K to 5	Rural	417	32%	23%	2%
22	K to 5	Suburban	1000	65%	26%	0%
Average			**563**	**49%**	**31%**	**2%**

Table 2
Demographic Characteristics of Florida *SAXON MATH* Grade 5 Schools Included in the Study

School	*Grades*	*Locale*	*Enrollment*	*% Free/Reduced Lunch*	*% Minority*	*% Limited English Proficient*
1	K to 5	Suburban	496	13%	23%	3%
2	K to 5	Suburban	714	18%	19%	3%
3	K to 5	Rural	417	32%	23%	2%
4	K to 5	Rural	536	33%	7%	2%
5	K to 5	Urban	897	34%	14%	1%
6	K to 5	Suburban	883	34%	9%	1%
7	K to 5	Suburban	494	35%	83%	16%
8	K to 5	Suburban	488	35%	26%	1%
9	K to 5	Suburban	414	41%	9%	1%
10	K to 5	Suburban	842	43%	10%	0%
11	K to 5	Suburban	485	45%	10%	1%
12	K to 5	Urban	433	67%	36%	0%
13	K to 5	Urban	290	85%	80%	0%
14	K to 5	Urban	436	89%	70%	1%
Average			**559**	**43%**	**30%**	**2%**

Description of the Florida Comprehensive Assessment Test (FCAT)

The following explanation of the FCAT was taken from the Florida State Department of Education Web site (Florida State Department of Education, 2009):

> ***What kind of test is the FCAT?***
> *The FCAT is made up of two kinds of tests: a criterion-referenced test (CRT), which measures how well students are meeting the* Sunshine State Standards *in reading, writing, mathematics, and science, and a norm-referenced test (NRT), which allows educators and parents to compare Florida student performance on reading and mathematics with the performance of students nationwide.*
>
> *Through a contract with a test publishing company, the Florida Department of Education developed FCAT Reading and Mathematics and first administered the test to students in Grades 4, 5, 8, and 10 in 1998. The FCAT was expanded to include Grades 3 through 10 in 2001 and to include FCAT Science in 2003*
>
> *The FCAT Reading, Mathematics, and Science tests require students to analyze, synthesize, and evaluate the information presented and to apply strategies or procedures they have learned. The level of thinking required of students goes beyond the recall of facts and literal comprehension required in many standardized tests.*

Tables 3 and 4 below indicate the grade 3 and grade 5 benchmarks from the *Sunshine State Standards* in Mathematics assessed by the FCAT.

Table 3
Grade 3 Benchmarks from the *Sunshine State Standards* in Mathematics Assessed on the Florida Comprehensive Assessment Test

Benchmark Number	Descriptor
MA.3.A.1.1	Model multiplication and division including problems presented in context: repeated addition, multiplicative comparison, array, how many combinations, measurement, and partitioning.
MA.3.A.1.2	Solve multiplication and division fact problems by using strategies that result from applying number properties.
MA.3.A.1.3	Identify, describe, and apply division and multiplication as inverse operations.
MA. 3.A.2.1	Represent fractions, including fractions greater than one, using area, set, and linear models.
MA.3.A.2.2	Describe how the size of the fractional part is related to the number of equal sized pieces in the whole.
MA.3.A.2.3	Compare and order fractions, including fractions greater than one, using models and strategies.
MA.3.A.2.4	Use models to represent equivalent fractions, including fractions greater than 1, and identify representations of equivalence.
MA.3.A.4.1	Create, analyze, and represent patterns and relationships using words, variables, tables, and graphs.
MA.3.A.6.1	Represent, compute, estimate, and solve problems using numbers through hundred thousands.
MA.3.A.6.2	Solve non-routine problems by making a table, chart, or list and searching for patterns.
MA.3.G.3.1	Describe, analyze, compare, and classify two-dimensional shapes using sides and angles - including acute, obtuse, and right angles - and connect these ideas to the definition of shapes.
MA.3.G.3.2	Compose, decompose, and transform polygons to make other polygons, including concave and convex polygons with three, four, five, six, eight, or ten sides.
MA.3.G.3.3	Build, draw, and analyze two-dimensional shapes from several orientations in order to examine and apply congruence and symmetry.
MA.3.G.5.1	Select appropriate units, strategies, and tools to solve problems involving perimeter.
MA.3.G.5.2	Measure objects using fractional parts of linear units such as 1/2, 1/4, and 1/10.
MA.3.G.5.3	Tell time to the nearest minute and to the nearest quarter hour, and determine the amount of time elapsed.
MA.3.S.7.1	Construct and analyze frequency tables, bar graphs, pictographs, and line plots from data, including data collected through observations, surveys, and experiments.

Table 4
Grade 5 Benchmarks from the *Sunshine State Standards* in Mathematics Assessed on the Florida Comprehensive Assessment Test

Benchmark Number	Descriptor
MA.5.A.1.1	Describe the process of finding quotients involving multi-digit dividends using models, place value, properties, and the relationship of division to multiplication.
MA.5.A.1.2	Estimate quotients or calculate them mentally depending on the context and numbers involved.
MA.5.A.1.3	Interpret solutions to division situations including those with remainders depending on the context of the problem.
MA.5.A.1.4	Divide multi-digit whole numbers fluently, including solving real-world problems, demonstrating understanding of the standard algorithm and checking the reasonableness of results.
MA.5.A.2.1	Represent addition and subtraction of decimals and fractions with like and unlike denominators using models, place value, or properties.
MA.5.A.2.2	Add and subtract fractions and decimals fluently, and verify the reasonableness of results, including in problem situations.
MA.5.A.2.3	Make reasonable estimates of fraction and decimal sums and differences, and use techniques for rounding.
MA.5.A.2.4	Determine the prime factorization of numbers.
MA.5.A.4.1	Use the properties of equality to solve numerical and real world situations.
MA.5.A.4.2	Construct and describe a graph showing continuous data, such as a graph of a quantity that changes over time.
MA.5.A.6.1	Identify and relate prime and composite numbers, factors, and multiples within the context of fractions.
MA.5.A.6.2	Use the order of operations to simplify expressions which include exponents and parentheses.
MA.5.A.6.3	Describe real-world situations using positive and negative numbers.
MA.5.A.6.4	Compare, order, and graph integers, including integers shown on a number line.
MA.5.A.6.5	Solve non-routine problems using various strategies including “solving a simpler problem” and “guess, check, and revise.”
MA.5.G.3.1	Analyze and compare the properties of two-dimensional figures and three-dimensional solids (polyhedra), including the number of edges, faces, vertices, and types of faces.
MA.5.G.3.2	Describe, define, and determine surface area and volume of prisms by using appropriate units and selecting strategies and tools.
MA.5.G.5.1	Identify and plot ordered pairs on the first quadrant of the coordinate plane.
MA.5.G.5.2	Compare, contrast, and convert units of measure within the same dimension (length, mass, or time) to solve problems.
MA.5.G.5.3	Solve problems requiring attention to approximation, selection of appropriate measuring tools, and precision of measurement.
MA.5.G.5.4	Derive and apply formulas for areas of parallelograms, triangles, and trapezoids from the area of a rectangle.
MA.5.S.7.1	Construct and analyze line graphs and double bar graphs.
MA.5.S.7.2	Differentiate between continuous and discrete data, and determine ways to represent those using graphs and diagrams.

Data Analyses

The mean standard score for each Florida *SAXON MATH* school from the spring 2005 and spring 2008 administrations of the mathematics portion of the FCAT were available from the FSDOE Web site. In addition, the FSDOE Web site provides the percentage of students at each school achieving at each of five performance levels:

Level 5: This student has success with the most challenging content of the *Sunshine State Standards*. A student scoring in Level 5 answers most of the test questions correctly, including the most challenging questions.

Level 4: This student has success with the challenging content of the *Sunshine State Standards*. A student scoring in Level 4 answers most of the test questions correctly, but may have only some success with questions that reflect the most challenging content.

Level 3: This student has partial success with the challenging content of the *Sunshine State Standards*, but performance is inconsistent. A student scoring in Level 3 answers many of the test questions correctly but is generally less successful with questions that are the most challenging.

Level 2: This student has limited success with the challenging content of the *Sunshine State Standards*.

Level 1: This student has little success with the challenging content of the *Sunshine State Standards*.

The standard scores were used to determine whether student performance on the math portion of the FCAT increased significantly from the spring 2005 test administration (pretest) to the spring 2008 test administration (posttest) for grade 3 and grade 5 students at Florida *SAXON MATH* schools included in the study. The following analyses were conducted:

- A Paired Comparison *t*-test was conducted to determine whether the pretest to posttest standard score gains of the total group of grade 3 and grade 5 students at Florida *SAXON MATH* schools were statistically significant.
- A Paired Comparison *t*-test was conducted to determine whether the pretest to posttest standard score gains of grade 3 and grade 5 students at both lower and higher socio-economic status Florida *SAXON MATH* schools were statistically significant.
- A Paired Comparison *t*-test was conducted to determine whether the pretest to posttest gains of grade 3 and grade 5 students at lower scoring pretest and higher scoring pretest Florida *SAXON MATH* schools (as determined by the average pretest scores of the total group of grade 3 and grade 5 students at each school) were statistically significant.

Grade 3 Pretest/Posttest Analyses of *SAXON MATH* Schools

Whole Group Pretest/Posttest Analyses

Researchers at ERIA conducted a Paired Comparison *t*-test to determine whether the pretest to posttest gains of the total group of grade 3 students at Florida *SAXON MATH* schools were statistically significant. The .05 level of significance was used as the level at which differences would be considered statistically significant. For the grade 3 analyses, 22 schools were included.

In addition to the Paired Comparison *t*-test, an effect-size analysis was computed. Cohen's *d* statistic was used to determine the effect size. This statistic provides an indication of the strength of the effect of the treatment regardless of the statistical significance. Cohen's *d* statistic is interpreted as follows:

.2 = small effect
.5 = medium effect
.8 = large effect

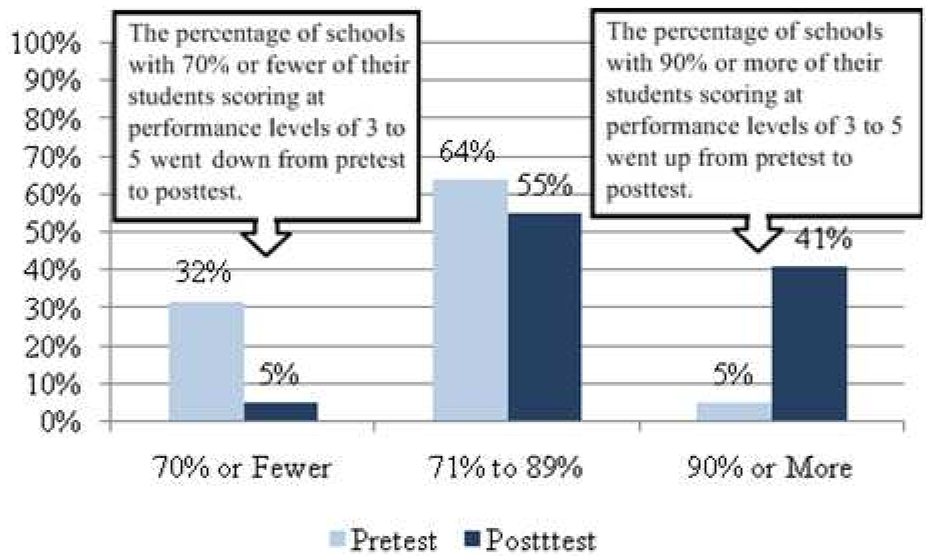
Table 5 presents the results of the *t*-test performed to determine if the pretest to posttest performance gains for grade 3 students at Florida *SAXON MATH* schools were statistically significant. The mean standard score was 326.8 on the pretest and 348.4 on the posttest, a difference that was statistically significant at the .0001 level. This level of significance indicates that such a difference would have occurred by chance less than once out of 10,000 repetitions. The effect size was large.

Table 5
Results Comparing the Average FCAT Math Standard Scores of Grade 3 Students at Florida *SAXON MATH* Schools in Spring 2005 (Pretest) and in Spring 2008 (Posttest)

Test	*Number of Schools*	*Mean Standard Score*	*SD*	*t*-Test	*Significance*	*Effect Size*
Pretest	22	326.8	16.7	6.540	<.0001	1.25
Posttest	22	348.4	17.6			

Figure 1 compares the percentage of Florida grade 3 *SAXON MATH* schools with various ranges of percentages of students scoring at performance levels of 3 to 5 (70% or fewer, 71% to 89% and 90% or more) at the time of the pretest and at the time of the posttest. The figure shows that the percentage of schools with 70% or fewer of their students scoring at levels 3 to 5 declined sharply while the percentage of schools with 90% or more of their students scoring at performance levels of 3 to 5 increased dramatically from pretest to posttest.

Figure 1
Percentage of Florida Grade 3 *SAXON MATH* Schools with Various Ranges of Percentages of Students Scoring at Performance Levels of 3 to 5 on the FCAT in Spring 2005 (Pretest) and in Spring 2008 (Posttest)



Socio-Economic Group Pretest/Posttest Analyses

A Paired Comparison *t*-test was used to compare the pretest and posttest scores of the grade 3 Florida *SAXON MATH* schools categorized as being of higher and lower socio-economic status (SES). The percentage of students receiving free and reduced lunch was used as the indicator of SES for this comparison. Schools were ranked from highest to lowest according to the percentage of students receiving free and reduced lunch at each school. That list was then divided in half with 11 schools in both the lower and higher free/reduced lunch groups. The percentage of students on free/reduced lunch programs at schools in the lowest half ranged from 19% to 45% with an average of 35%. This group was considered the higher SES group of schools since they had the fewest students enrolled in free/reduced lunch programs. The percentage of students on free/reduced lunch programs at schools in the highest half ranged from 45% to 89% with an average of 63%. This group was considered the lower SES group of schools since they had the most students enrolled in free/reduced lunch programs.

The .05 level of significance was used as the level at which increases would be considered statistically significant.

Table 6 presents the results of the *t*-test performed to determine if the pretest to posttest standard score gains of grade 3 students at lower and higher SES Florida *SAXON MATH* schools were statistically significant. For the lower SES schools, the mean standard score on the pretest was 320.5 and on the posttest the mean standard score was 340.8, a difference that was statistically significant at the .002 level. This level of significance indicates that such a difference would have occurred by chance less two times out of 1,000 repetitions. The effect size was large.

For the higher SES schools, the mean standard score on the pretest was 333.0 and on the posttest the mean standard score was 355.9, a difference that was statistically significant at the .001 level. This level of significance indicates that such a difference would have occurred by chance less one time out of 1,000 repetitions. The effect size was large.

While students at the lower SES schools started out with a lower mean standard score than the students at the higher SES schools, the two groups made about the same standard score increase and the increases for both groups were statistically significant.

Table 6
Results Comparing the Mean FCAT Mathematics Standard Scores of Grade 3 Students at Florida *SAXON MATH* Schools in Spring 2005 (Pretest) and in Spring 2008 (Posttest) For High and Low SES Schools

Test	***Number of Schools***	***Mean Standard Score***	***SD***	***t*-Test**	***Significance***	***Effect Size***
Lower Socio-Economic Schools						
Pretest	11	320.5	18.0	4.183	<.002	1.03
Posttest	11	340.8	21.0			
Higher Socio-Economic Schools						
Pretest	11	333.0	13.3	4.891	<.001	2.00
Posttest	11	355.9	9.2			

Figure 2 compares the percentage of Florida grade 3 *SAXON MATH* lower SES schools with various ranges of percentages of students scoring at performance levels of 3 to 5 (70% or fewer, 71% to 89% and 90% or more) at the time of the pretest and at the time of the posttest. The figure shows that the percentage of schools with 70% or fewer of their students scoring at levels 3 to 5 declined sharply while the percentage of schools with 90% or more of their students scoring at performance levels of 3 to 5 increased from zero to 18%.

Figure 2
Percentage of Florida Grade 3 *SAXON MATH* Lower SES Schools with Various Ranges of Percentages of Students Scoring at Performance Levels of 3 to 5 on the FCAT in Spring 2005 (Pretest) and in Spring 2008 (Posttest)

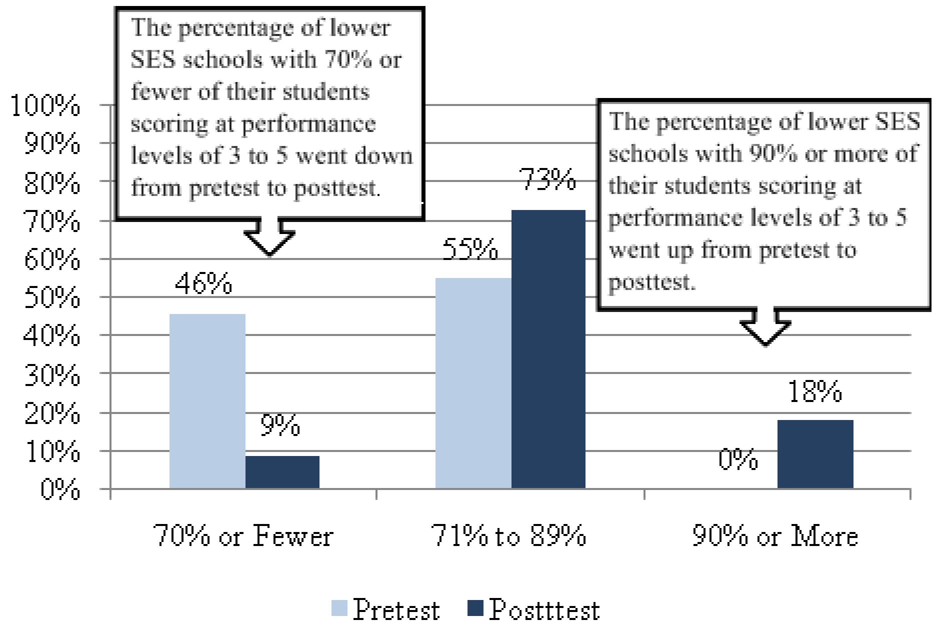
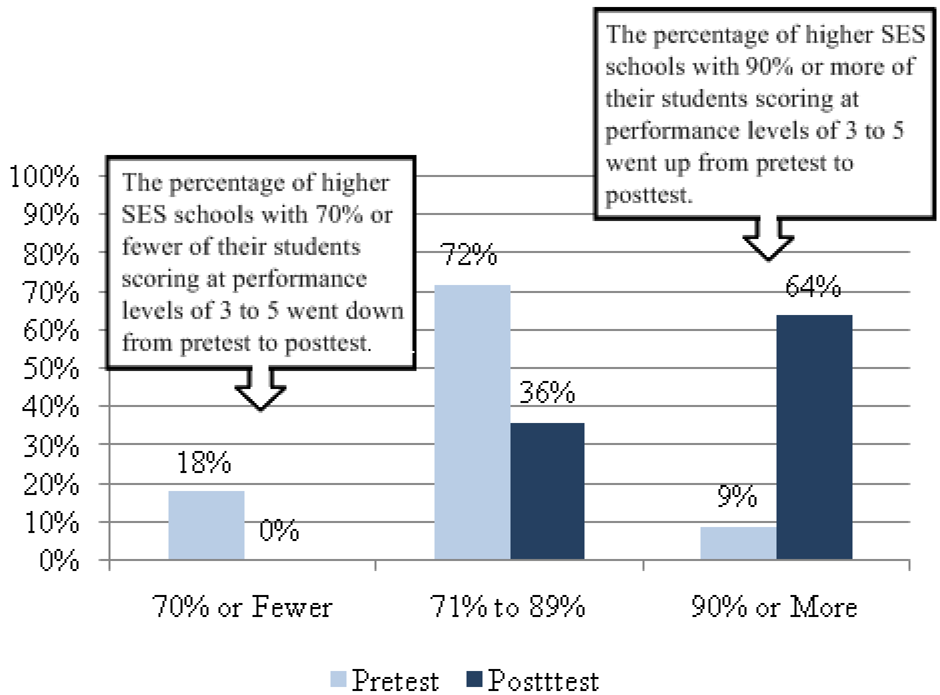


Figure 3 compares the percentage of Florida grade 3 *SAXON MATH* higher SES schools with various ranges of percentages of students scoring at performance levels of 3 to 5 (70% or fewer, 71% to 89% and 90% or more) at the time of the pretest and at the time of the posttest. The figure shows that the percentage of schools with 70% or fewer of their students scoring at levels 3 to 5 declined from 18% to zero while the percentage of schools with 90% or more of their students scoring at performance levels of 3 to 5 increased dramatically from pretest to posttest.

Figure 3
Percentage of Florida Grade 3 *SAXON MATH* Higher SES Schools with Various Ranges of Percentages of Students Scoring at Performance Levels of 3 to 5 on the FCAT in Spring 2005 (Pretest) and in Spring 2008 (Posttest)



Pretest Score Group Pretest/Posttest Analyses

The grade 3 schools were divided into two approximately equal groups based on the average pretest score of the total group of grade 3 students at each school. Each group included 11 schools. Paired Comparison *t*-tests were conducted to determine if both groups made significant pretest to posttest gains.

Table 7 presents the results of the *t*-tests performed to determine if the pretest to posttest gains of grade 3 students at Florida *SAXON MATH* schools in both the lower and higher scoring pretest groups were statistically significant. The mean standard score for the lower scoring group increased from 314.5 to 340.8. The difference for the lower scoring pretest group was statistically significant at the .0001 level, indicating a change that would have occurred by chance less than once time out of 10,000 repetitions. The effect size was large.

The mean standard score for the higher scoring group increased from 339.0 to 355.9. The difference for the higher scoring pretest group was statistically significant at the .002 level, indicating a change that would have occurred by chance less than twice out of 1,000 repetitions. The effect size was large.

Table 7

Results Comparing the FCAT Math Standard Scores of Grade 3 Students at Florida *SAXON MATH* Schools in Spring 2005 (Pretest) and in Spring 2008 (Posttest) For Lower and Higher Scoring Pretest Groups

Test	*Number of Schools*	*Mean Standard Score*	*SD*	t-*test*	*Significance*	*Effect Size*
Lower Scoring Pretest Schools						
Pretest	11	314.5	14.2	5.143	<.0001	1.56
Posttest	11	340.8	19.2			
Higher Scoring Pretest Schools						
Pretest	11	339.0	7.3	4.323	<.002	4.323
Posttest	11	355.9	12.6			

Figure 4 compares the percentage of Florida grade 3 *SAXON MATH* lower pretest schools with various ranges of percentages of students scoring at performance levels of 3 to 5 (70% or fewer, 71% to 89% and 90% or more) at the time of the pretest and at the time of the posttest. The figure shows that the percentage of schools with 70% or fewer of their students scoring at levels 3 to 5 declined sharply while the percentage of schools with 90% or more of their students scoring at performance levels of 3 to 5 increased from zero to 18%.

Figure 4
Percentage of Florida Grade 3 *SAXON MATH* Lower Pretest Schools with Various Ranges of Percentages of Students Scoring at Performance Levels of 3 to 5 on the FCAT in Spring 2005 (Pretest) and in Spring 2008 (Posttest)

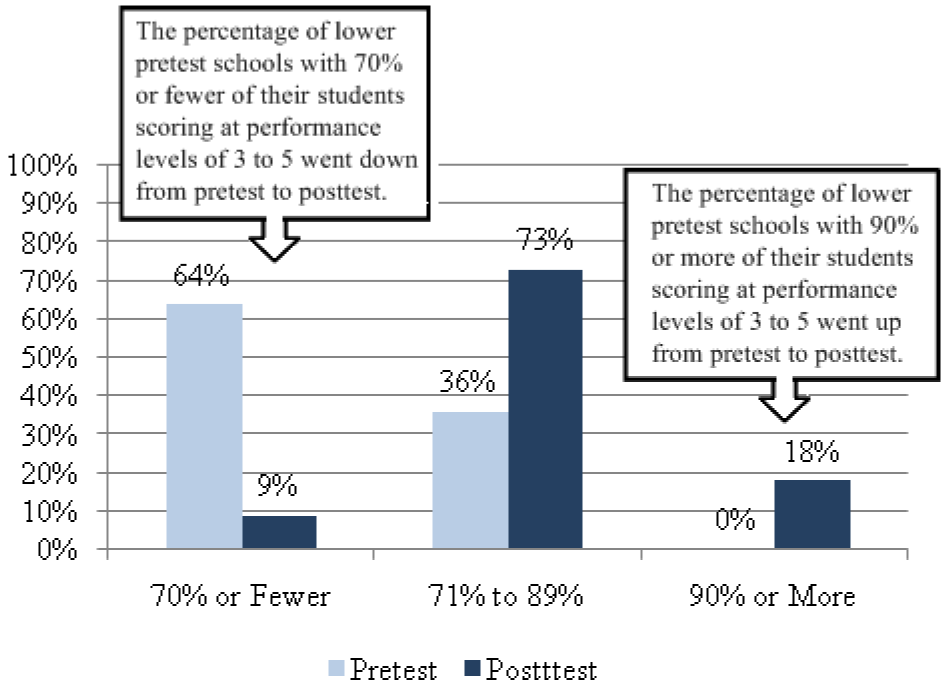
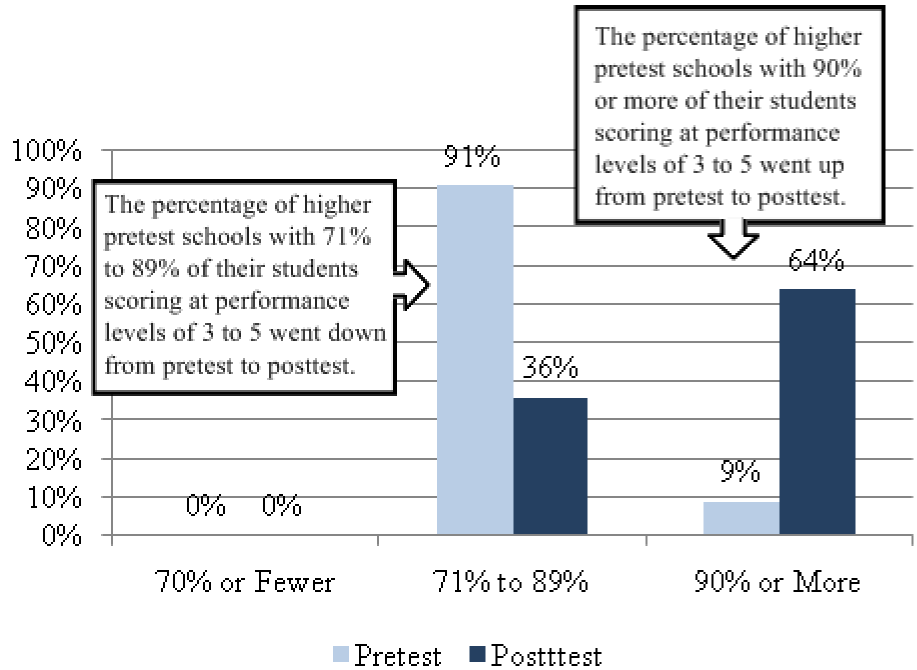


Figure 5 compares the percentage of Florida grade 3 *SAXON MATH* higher pretest schools with various ranges of percentages of students scoring at performance levels of 3 to 5 (70% or fewer, 71% to 89% and 90% or more) at the time of the pretest and at the time of the posttest. The figure shows that the there were no higher pretest schools with 70% or fewer of their students scoring at levels 3 to 5 at the time of the pretest or the posttest. The percentage of higher pretest schools with 90% or more of their students scoring at performance levels of 3 to 5 increased from 9% to 64% from pretest to posttest.

Figure 5
Percentage of Florida Grade 3 *SAXON MATH* Higher Pretest Schools with Various Ranges of Percentages of Students Scoring at Performance Levels of 3 to 5 on the FCAT in Spring 2005 (Pretest) and in Spring 2008 (Posttest)



Grade 5 Pretest/Posttest Analyses of *SAXON MATH* Schools

Whole Group Pretest/Posttest Analyses

Researchers at ERIA conducted a Paired Comparison *t*-test to determine whether the pretest to posttest gains of the total group of grade 5 students at Florida *SAXON MATH* schools were statistically significant. The .05 level of significance was used as the level at which differences would be considered statistically significant. For the grade 5 analyses, 14 schools were included.

In addition to the Paired Comparison *t*-test, an effect-size analysis was computed. Cohen's *d* statistic was used to determine the effect size. This statistic provides an indication of the strength of the effect of the treatment regardless of the statistical significance. Cohen's *d* statistic is interpreted as follows:

.2 = small effect
.5 = medium effect
.8 = large effect

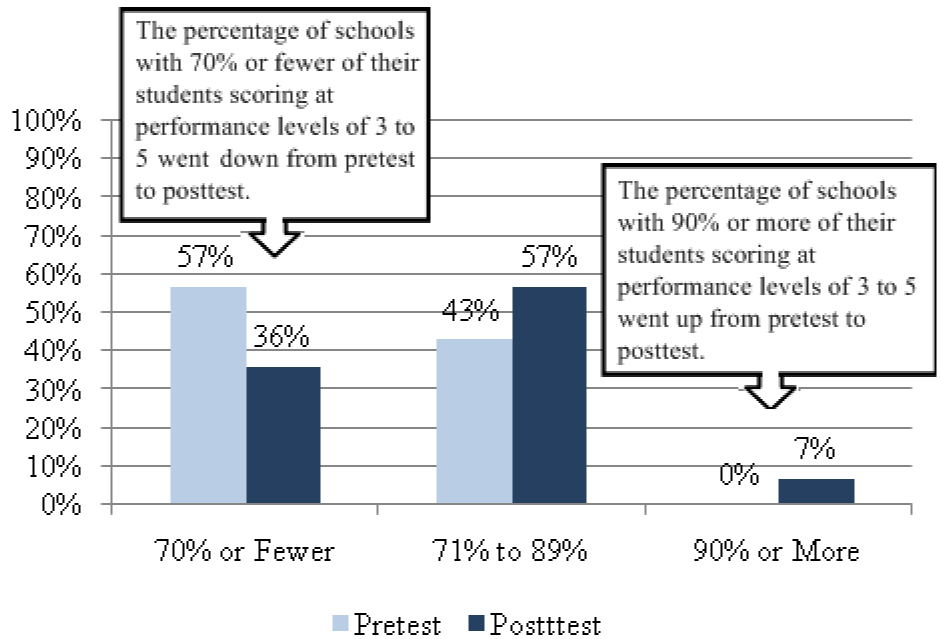
Table 8 presents the results of the *t*-test performed to determine if the pretests to posttest performance gains for grade 5 students at Florida *SAXON MATH* schools were statistically significant. The mean standard score was 327.9 on the pretest and 351.5 on the posttest, a difference that was statistically significant at the .005 level. This level of significance indicates that such a difference would have occurred by chance less than five times out of 1,000 repetitions. The effect size was large.

Table 8
Results Comparing the Average FCAT Math Standard Scores of Grade 5 Students at Florida *SAXON MATH* Schools in Spring 2005 (Pretest) and in Spring 2008 (Posttest)

Test	*Number of Schools*	*Mean Standard Score*	*SD*	*t*-Test	*Significance*	*Effect Size*
Pretest	14	327.9	23.6	3.368	<.005	1.21
Posttest	14	351.5	14.4			

Figure 6 compares the percentage of Florida grade 5 *SAXON MATH* schools with various ranges of percentages of students scoring at performance levels of 3 to 5 (70% or fewer, 71% to 89% and 90% or more) at the time of the pretest and at the time of the posttest. The figure shows that the percentage of schools with 70% or fewer of their students scoring at levels 3 to 5 declined while the percentage of schools with 90% or more of their students scoring at performance levels of 3 to 5 increased from pretest to posttest.

Figure 6
Percentage of Florida Grade 5 *SAXON MATH* Schools with Various Ranges of Percentages of Students Scoring at Performance Levels of 3 to 5 on the FCAT in Spring 2005 (Pretest) and in Spring 2008 (Posttest)



Socio-Economic Group Pretest/Posttest Analyses
A Paired Comparison *t*-test was used to compare the pretest and posttest scores of the grade 5 Florida *SAXON MATH* schools categorized as being of higher and lower socio-economic status (SES). The percentage of students receiving free and reduced lunch was used as the indicator of SES for this comparison. Schools were ranked from highest to lowest according to the percentage of students receiving free and reduced lunch at each school. That list was then divided in half with 7 schools in both the lower and higher free/reduced lunch groups. The percentage of students on free/reduced lunch programs at schools in the lowest half ranged from 13% to 35% with an average of 28%. This group was considered the higher SES group of schools since they had the fewest students enrolled in free/reduced lunch programs. The percentage of students on free/reduced lunch programs at schools in the highest half ranged from 35% to 89% with an average of 58%. This group was considered the lower SES group of schools since they had the most students enrolled in free/reduced lunch programs.

The .05 level of significance was used as the level at which increases would be considered statistically significant.

Table 9 presents the results of the *t*-test performed to determine if the pretest to posttest standard score gains of grade 5 students at lower and higher SES Florida *SAXON MATH* schools were statistically significant. For the lower SES schools, the mean standard score on the pretest was 340.9 and on the posttest the mean standard score was 351.4, a difference that was not statistically significant. The effect size was large. The reason for the large effect size in the absence of a statistically significant difference between the pretest and posttest scores is due, in large part, to the very small sample size. The statistical significance is affected by sample size whereas the effect size is not.

For the higher SES schools, the mean standard score on the pretest was 314.9 and on the posttest the mean standard score was 351.6, a difference that was statistically significant at the .01 level. This level of significance indicates that such a difference would have occurred by chance less one time out of 100 repetitions. The effect size was large.

Table 9
Results Comparing the Mean FCAT Mathematics Standard Scores of Grade 5 Students at Florida *SAXON MATH* Schools in Spring 2005 (Pretest) and in Spring 2008 (Posttest) For High and Low SES Schools

Test	*Number of Schools*	*Mean Standard Score*	*SD*	*t-Test*	*Significance*	*Effect Size*
Lower Socio-Economic Schools						
Pretest	7	340.9	12.3	1.554	Non-Sig.	.81
Posttest	7	351.4	13.5			
Higher Socio-Economic Schools						
Pretest	11	314.9	25.7	3.496	<.01	1.7
Posttest	11	351.6	16.4			

Figure 7 compares the percentage of Florida grade 5 *SAXON MATH* lower SES schools with various ranges of percentages of students scoring at performance levels of 3 to 5 (70% or fewer, 71% to 89% and 90% or more) at the time of the pretest and at the time of the posttest. The figure shows that the percentage of schools with 70% or fewer of their students scoring at levels 3 to 5 declined from pretest to posttest. While there were no schools at the time of the pretest or the posttest with 90% or more of their students scoring at performance levels of 3 to 5, the percentage of schools with 71% to 89% of their students scoring at performance levels of 3 to 5 increased from 57% to 71% from pretest to posttest.

Figure 7
Percentage of Florida Grade 5 *SAXON MATH* Lower SES Schools with Various Ranges of Percentages of Students Scoring at Performance Levels of 3 to 5 on the FCAT in Spring 2005 (Pretest) and in Spring 2008 (Posttest)

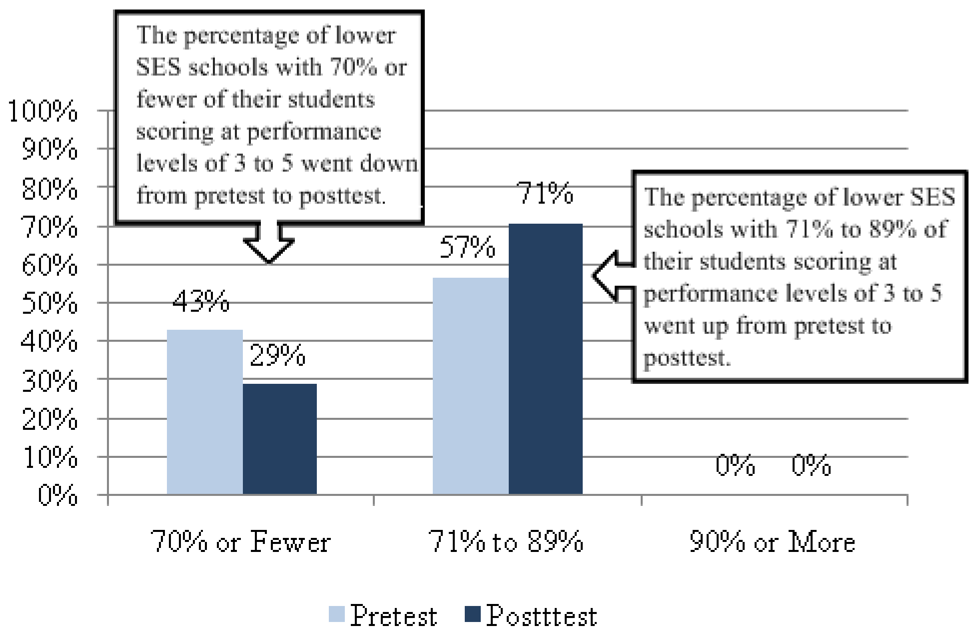
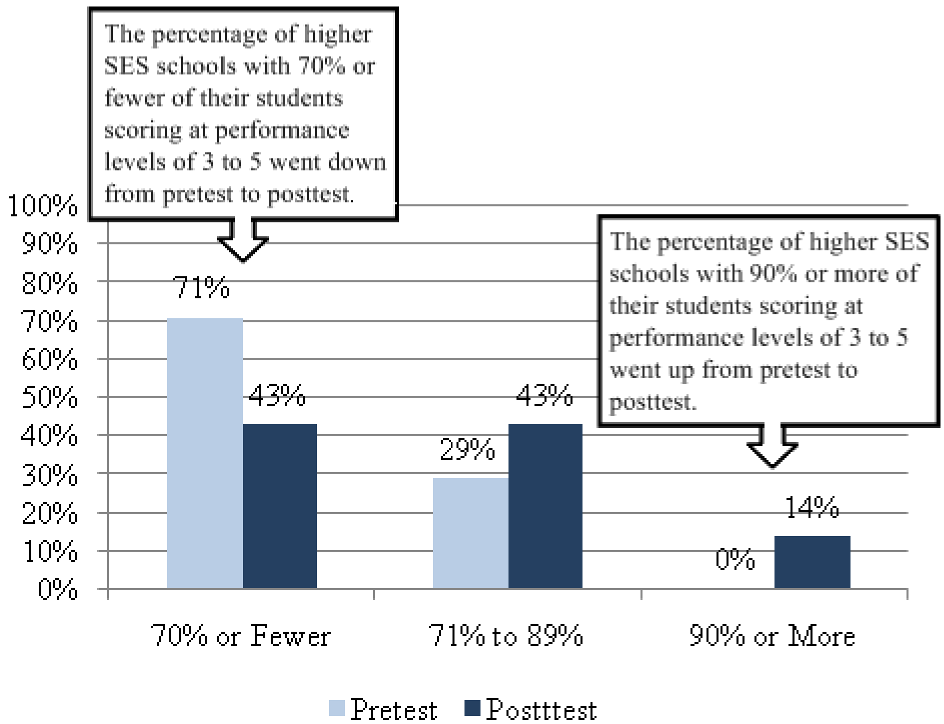


Figure 8 compares the percentage of Florida grade 5 *SAXON MATH* higher SES schools with various ranges of percentages of students scoring at performance levels of 3 to 5 (70% or fewer, 71% to 89% and 90% or more) at the time of the pretest and at the time of the posttest. The figure shows that the percentage of schools with 70% or fewer of their students scoring at levels 3 to 5 declined from 71% to 43% while the percentage of schools with 90% or more of their students scoring at performance levels of 3 to 5 increased from zero to 14% from pretest to posttest.

Figure 8
Percentage of Florida Grade 5 *SAXON MATH* Higher SES Schools with Various Ranges of Percentages of Students Scoring at Performance Levels of 3 to 5 on the FCAT in Spring 2005 (Pretest) and in Spring 2008 (Posttest)



Pretest Score Group Pretest/Posttest Analyses

The grade 5 schools were divided into two approximately equal groups based on the average pretest score of the total group of grade 5 students at each school. Each group included 7 schools. Paired Comparison *t*-tests were conducted to determine if both groups made significant pretest to posttest gains.

Table 10 presents the results of the *t*-test performed to determine if the pretest to posttest gains of grade 5 students at Florida *SAXON MATH* schools in both the lower and higher scoring pretest groups were statistically significant. The mean standard score for the lower scoring group increased from 309.4 to 351.0. The difference for the lower scoring pretest group was statistically significant at the .004 level, indicating a change that would have occurred by chance less than four out of 1,000 repetitions. The effect size was large.

The mean standard score for the higher scoring group increased from 346.2 to 352.0. The difference for the higher scoring pretest group was not statistically significant. The effect size was medium.

Table 10
Results Comparing the FCAT Math Standard Scores of Grade 5 Students at Florida *SAXON MATH* Schools in Spring 2005 (Pretest) and in Spring 2008 (Posttest) For Lower and Higher Scoring Pretest Groups

Test	*Number of Schools*	*Mean Standard Score*	*SD*	t-*test*	*Significance*	*Effect Size*
Lower Scoring Pretest Schools						
Pretest	7	309.4	18.8	4.590	<.004	2.38
Posttest	7	351.0	16.0			
Higher Scoring Pretest Schools						
Pretest	7	346.2	7.6	1.159	Non-Sig.	.50
Posttest	7	352.0	14.0			

Figure 9 compares the percentage of Florida grade 5 *SAXON MATH* lower pretest schools with various ranges of percentages of students scoring at performance levels of 3 to 5 (70% or fewer, 71% to 89% and 90% or more) at the time of the pretest and at the time of the posttest. The figure shows that all of the schools had 70% or fewer of their students scoring at levels 3 to 5 at the time of the pretest. However, at the time of the posttest, 57% of the schools had 71% to 89% of their students scoring at performance levels of 3 to 5.

Figure 9
Percentage of Florida Grade 5 *SAXON MATH* Lower Pretest Schools with Various Ranges of Percentages of Students Scoring at Performance Levels of 3 to 5 on the FCAT in Spring 2005 (Pretest) and in Spring 2008 (Posttest)

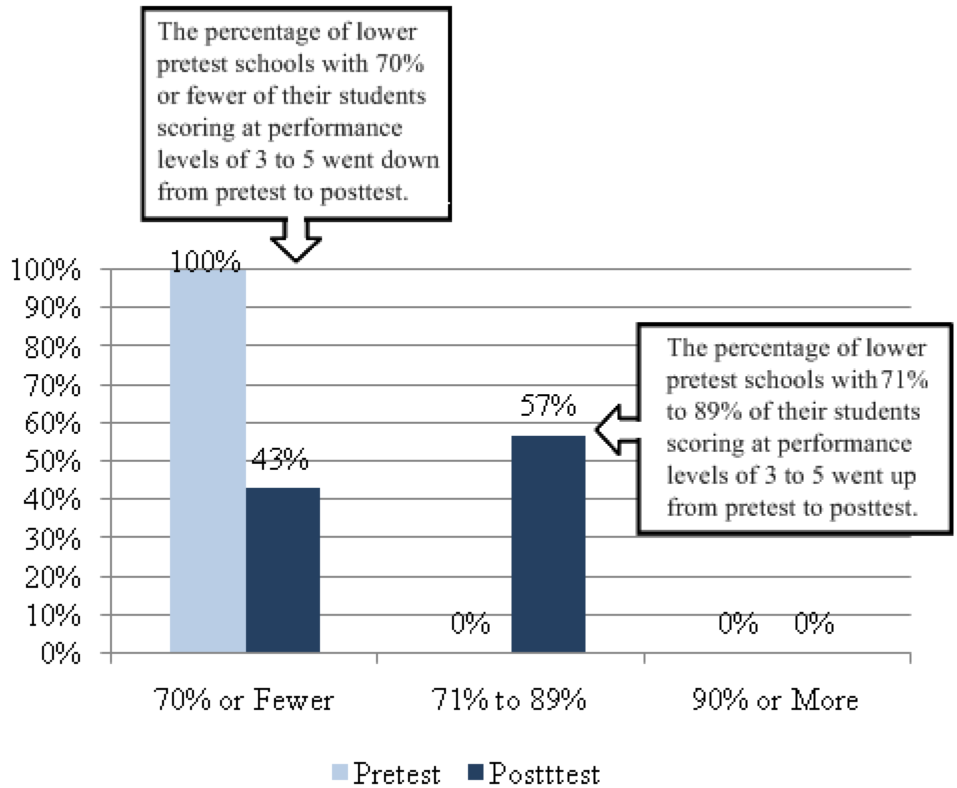
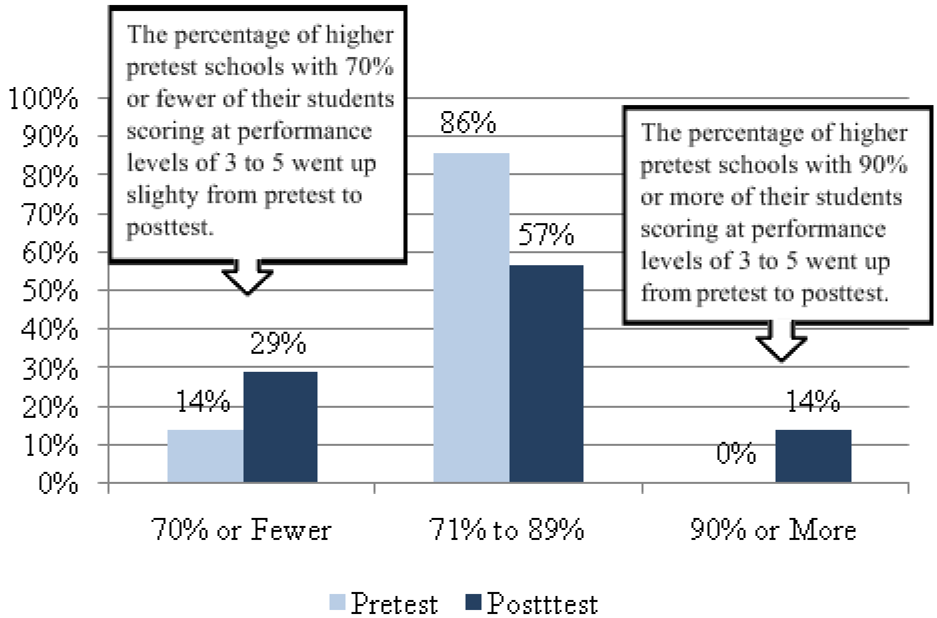


Figure 10 compares the percentage of Florida grade 5 *SAXON MATH* higher pretest schools with various ranges of percentages of students scoring at performance levels of 3 to 5 (70% or fewer, 71% to 89% and 90% or more) at the time of the pretest and at the time of the posttest. The figure shows that the percentage of schools with 90% or more of their students scoring at performance levels of 3 to 5 increased from zero to 14%.

Figure 10
Percentage of Florida Grade 5 *SAXON MATH* Higher Pretest Schools with Various Ranges of Percentages of Students Scoring at Performance Levels of 3 to 5 on the FCAT in Spring 2005 (Pretest) and in Spring 2008 (Posttest)



Conclusions

This study sought to determine the effect of the *SAXON MATH* program on students' math skills and strategy use.

When comparing the pretest to posttest gains made by grade 3 and grade 5 students at Florida *SAXON MATH* schools, gains were statistically significant at both grade levels for the total groups. In addition, significant gains were made by grade 3 students at lower and at higher SES *SAXON MATH* schools and by grade 5 students at higher SES SAXON *MATH* schools. When Florida *SAXON MATH* schools were divided into lower pretest schools and higher pretest schools based on the pretest scores of the total group, the lower pretest groups at both grade levels and the higher pretest group at grade 3 made significant pretest to posttest gains. A summary of the results is provided in Table 11 below. The table indicates whether the gains were significant as well as the effect size of each significant gain.

Table 11
Summary of the Pretest/Posttest Score Analyses Conducted to Determine if Significant Gains were Made on the Math Portion of the FCAT for Grade 3 and Grade 5 Students at Florida *SAXON MATH* Schools

	Grade 3		***Grade 5***	
Group	***Gain Statistically Significant?***	***Effect Size***	***Gain Statistically Significant?***	***Effect Size***
All *SAXON MATH* Schools	Yes	Large	Yes	Large
Lower SES *SAXON MATH* Schools	Yes	Large	No	Large
Higher SES *SAXON MATH* Schools	Yes	Large	Yes	Large
Lower Pretest Group *SAXON MATH* Schools	Yes	Large	Yes	Large
Higher Pretest Group *SAXON MATH* Schools	Yes	Large	No	Medium

This study sought to determine if *SAXON MATH* is instructionally effective. Based on the results of this study, instruction based on *SAXON MATH* significantly increases grade 3 and grade 5 students' knowledge and understanding of mathematics over a three year period in Florida schools using the *SAXON MATH* program.

References

Denton, K., & West, J. (2002). *Children's reading and mathematics achievement in kindergarten and first grade*. Washington, DC: National Center for Education Statistics.

Florida State Department of Education. (2009). *Florida Comprehensive Assessment Test.* Retrieved April 23, 2009, from the Florida State Department of Education Web site: http://fcat.fldoe.org/

Geary, D. C., Hoard, M.K., Byrd-Craven, J., & DeSoto, M.C. (2005). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Experimental Child Psychology, 88*, 121 - 151.

International Center for Leadership in Education. (2006). *Florida Curriculum Matrix Summary*. Retrieved March 16, 2009, from the International Center for Leadership in Education Web site: http://www.daggett.com/CM%20pdf%20files/OKICLEsummary2.pdf

Ma, X. (2000). A longitudinal assessment of antecedent course work in mathematics and subsequent mathematical attainment. *Journal of Educational Research, 94*, 16 - 28.

Magnuson, K., Myers, M., Ruhm, C., & Waldfogel, J. (2003). *Inequality in preschool education and school readiness*. New York, NY: Columbia University Press.

Mullis, I.V.S., Martin, M.O., Gonzalez E.J., & Chrostowki, S.J. (2005). *TIMSS 2003 International mathematics report: Findings from IEA's Trends in International Mathematics and Science study at the fourth and eighth grades*. Chestnut Hill, MA: Boston College, Center for the Study of Testing, Evaluation, and Educational Policy.

National Assessment of Educational Progress. (2007). *The nation's report card: Mathematics 2007*. Retrieved February 11, 2009, from the National Assessment of Educational Progress Web site: http://nces.ed.gov/nationsreportcard/pubs/main2007/2007494.asp

National Association for the Education of Young Children and National Council of Teachers of Mathematics. (2002). *Position statement. Early childhood mathematics: Promoting good beginnings*. Retrieved February 11, 2009, from the National Association for the Education of Young Children Web site: http://www.naeyc.org/about/positions/psmath.asp

National Council of Teachers of Mathematics. (2009). *Principals and standards for school mathematics*. Retrieved February 11, 2009, from the National Council of Teachers of Mathematics Web site: http://standards.nctm.org/document/chapter1/index.htm

Ostad, S. A. (1997). Developmental differences in addition strategies: A comparison of mathematically disabled and mathematically normal children. *British Journal of Educational Psychology*, *67*, 345 - 357.

Ostad, S. A. (1998). Developmental differences in solving simple arithmetic word problems and simple number-fact problems: A comparison of mathematically normal and mathematically disabled children. *Mathematical Cognition*, *4*(1), 1 - 19.

U.S. Department of Education. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel.* Retrieved February 11, 2009, from the U.S. Department of Education Web site: http://www.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf