Saxon Math: Summaries of Independent Studies

March 2005

This page intentionally left blank.

Saxon Math:
Summaries of Independent Studies

Table of Contents

Introduction

Because of the popularity and uniqueness of the *Saxon Math* program, the pedagogy and effectiveness of the program have been extensively and independently studied in universities, schools, and by various organizations. The recent demand to provide research-based products has created an increase in the need for product-specific research. Independently researched interventions provide the most unbiased source of efficacy information and achievement results.

In order to examine the available independent research from the research literature, Saxon Publishers asked a team at the University of Oklahoma to conduct a thorough literature search to locate evaluations that included *Saxon Math* products as part of a research-defined curriculum. While not an exhaustive compendium of all the independent research evaluating *Saxon Math*, the following document provides summaries of 14 independent studies covering a variety of *Saxon Math* products and grade levels. Taken together as a consensus of research, this document addresses the positive impact *Saxon Math* has had on the mathematical achievement of students.

The report is organized first by grade level and then by reverse chronology when more than one representation of a grade level is included. Use the following table of contents to find a research study by grade span or product (note that not all products have been individually evaluated).

By Grade Span:

By Product:

Elementary School

Second and Third Grade – *Math K-3*

Calvery, R., Bell, D., & Wheeler, G. (1993, November). *A comparison of selected second- and third-graders' math achievement: Saxon vs. Holt.* Paper presented at the annual meeting of the Mid-South Educational Research Association, New Orleans, Louisiana.

- This study presents the findings from a year-long comparison of *Saxon Math K-3* and the *Holt Math Series* at the second and third grade.

- Participants included 192 second- and third-grade students from a school located in rural Arkansas. Four second-grade and four third-grade classrooms of approximately 24 students each were evaluated for math achievement growth over the 1992-1993 school year. In both the second and third grades one of the classrooms implemented the Saxon program while the other three implemented the Holt program.

- Math achievement over time was measured using the Stanford Achievement Test (Eighth Edition, Primary) (SAT 8) as a pre- and posttest measure.

- When both second- and third-grade students were combined, the pretest results indicated that the students using the Saxon program scored significantly below the students using the Holt program. However, after a year of implementing the Saxon program, these students made statistically significantly higher achievement gains on the SAT 8 than the students using the Holt program. Despite starting at a lower achievement level, the Saxon program helped these students close the achievement gap and reach the level of higher performing students.[1]

- This same pattern was seen in both the second and third grades as well. It was found that the

---

[1] Note that graphical representations of the data for this study could not be created because the information in the provided tables seems to have been mistakenly entered and there is no method to verify or obtain the correct data. However, the text description of the findings in the report are believed to be accurate.

second-grade students using the Saxon program began the year at an even larger disadvantage but were still able to make statistically significantly higher gains on the posttest than the students using the Holt program.

Fourth Grade – *Math 4* and *Math 5/4*

Hansen, E., & Greene, K. (2000). *A recipe for math: What's cooking in the classroom?* Retrieved March 3, 2001, from http://www.secondaryenglish.com/recipeformath.html.

- This was a teacher-driven research study conducted by two teachers piloting the Saxon program in fourth-grade classrooms at their respective rural Georgia schools.

- At one of the participating schools, *Saxon Math 4*, the consumable fourth-grade product, was compared to the MacMillan program, *Math in Action* in two fourth-grade classrooms. These classes were taught by different teachers. At the second school, a cohort analysis was used to compare the achievement of two classes using the Saxon approach (one including advanced math students and the other including average math students) to the previous year's fourth-grade math class before the introduction of *Saxon Math*. The advanced group had the same teacher from year to year while the average group had different teachers. The classrooms piloting the Saxon program at this school used the *Saxon Math 5/4* book, the non-consumable version of the fourth-grade text, while the year before average group cohort used the *Holt Mathematics Unlimited* book and the advanced group used the Harcourt Brace basal.

- Grade equivalent scores from the Iowa Tests of Basic Skills (ITBS) were gathered to measure student achievement for all comparisons. ITBS scores from the previous year were used as the pretest measure, while ITBS scores of the implementation year were used as the posttest.

- Because of the informality of the report, the researchers did not provide all of the findings or data, therefore no graphical representations of the data could be provided.

- At the school using the *Saxon Math 4* consumable book, the Saxon group had lower

ITBS pretest scores and slightly higher posttest scores than the comparison group. While both groups experienced growth throughout the year, it was not determined if the growth seen in the Saxon group was significantly higher. However, the Saxon group made larger gains over time overall than the comparison group.
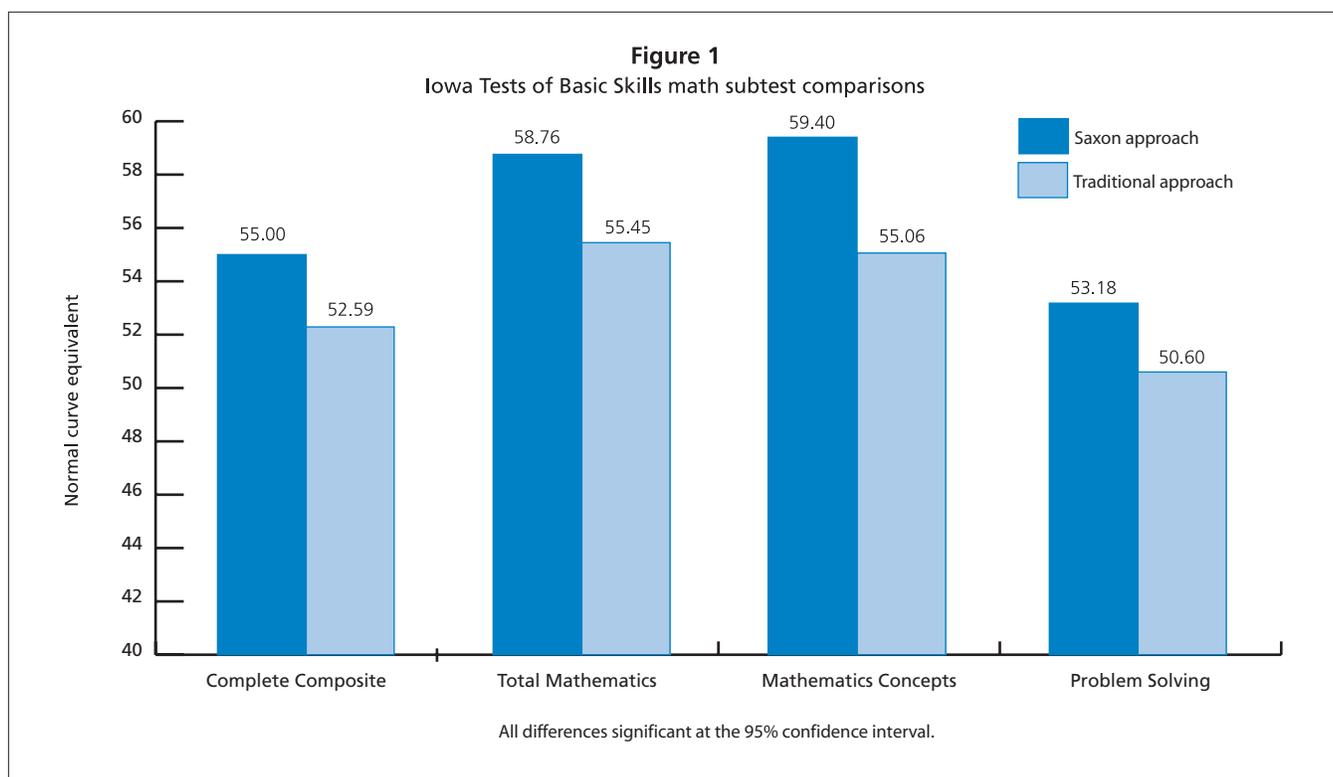
- For the school using the *Saxon Math 5/4* non-consumable book, only limited information is reported. The researchers indicate that the average mathematics students from the Saxon group had initially lower ITBS pretest scores but finished with higher scores on the posttest than the cohort comparison group. While determination of statistical significance is not possible, it is stated that the Saxon group grew 1.49 grade-equivalent levels, which is higher than the expected growth of 1.0 after a year-long implementation.

Grades K-5 – *Math K-5*

Nguyen, K., (1994). *The 1993-1994 Saxon mathematics evaluation report*. Oklahoma City, Oklahoma: Oklahoma City Public Schools.

- To evaluate the effectiveness of the *Saxon Math* elementary products, the researchers from the Oklahoma City Public Schools Research, Planning, and Evaluation Department examined student achievement from five Oklahoma City schools that fully implemented the *Saxon Math* program in grades K-5.[2] The schools participating in this study had been implementing the *Saxon Math* program for two years at the time of data collection.

- Student achievement on the subtests of the ITBS at the five Saxon schools was compared to a matched-sample of the students using a Scott Foresman text that was selected to be the control group. These students were matched to students using the Saxon program on grade level, gender, race, socio-economic status (SES), and the year prior ITBS total math score.

- The students using the *Saxon Math* program scored significantly higher than the control group on five out of the nine subtests of the ITBS: Complete Composite, Total Mathematics,

---

[2] The number of students participating in the evaluation could not be determined by the information provided in the text.

Figure 1
Iowa Tests of Basic Skills math subtest comparisons

Legend: Saxon approach / Traditional approach

Normal curve equivalent (y-axis: 40 to 60)

Subtest	Saxon approach	Traditional approach
Complete Composite	55.00	52.59
Total Mathematics	58.76	55.45
Mathematics Concepts	59.40	55.06
Problem Solving	53.18	50.60

All differences significant at the 95% confidence interval.

Mathematics Concepts, Problem Solving, and Reading Comprehension. The results pertaining to the math subtests can be found in Figure 1.[3]
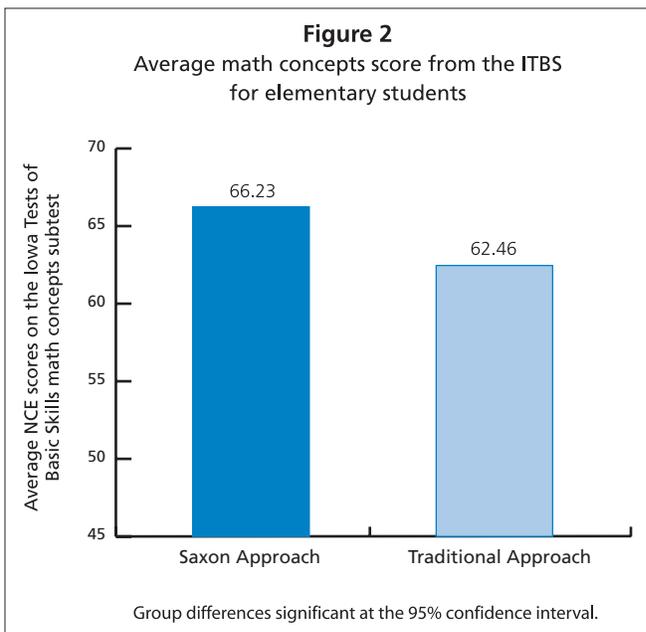
- Student achievement was also examined by grade level. Grades 3, 4, and 5 from the five Saxon schools and a matched control sample from non-Saxon schools were chosen for comparison. The Saxon group had higher achievement on 23 out of the 27 grade level comparisons on the ITBS subtests. Eleven of these differences were significant in favor of the Saxon group. There was not a specific pattern of results across grade levels, but in general, on the majority of grade level subtest comparisons, the Saxon group outperformed the control group.

Grades K-5 – *Math K-5*

Nguyen, K., & Elam, P. (1993). *The 1992-1993 Saxon mathematics evaluation report.* Oklahoma City, Oklahoma: Oklahoma City Public Schools.

- In order to evaluate the effectiveness of *Saxon Math* at the elementary level, a study was conducted by the Oklahoma City Public Schools, Planning, Research, and Evaluation Department to compare student achievement in math after using the *Saxon Math* program or a Scott Foresman program in Oklahoma City schools.[4]
- To achieve the most accurate results, a matched-sample of the students using the Scott Foresman text was selected to be the control group. These students were matched to students using the Saxon program on grade level, gender, race, SES, and the year prior ITBS total math score.
- The ITBS scores from the implementation year were collected and used to evaluate growth in mathematic skill over the implementation time for both groups. Because students between groups were matched on ITBS score, this equated the groups on math ability removing the need to use the pretest scores as a covariate in the analysis. On the posttest ITBS, the students using the *Saxon Math* products outscored the control group students on all subtests: Complete Composite, Total Math, Problem Solving, Reading Comprehension, Math Computation, Math Concepts, Science, and Social Studies. However, only the differences between groups

on the Math Concepts, Science, and Social Studies tests were significant (see Figure 2 for a graphical display of the math concepts subtest).[5]

Figure 2
Average math concepts score from the ITBS for elementary students

Saxon Approach: 66.23
Traditional Approach: 62.46

Average NCE scores on the Iowa Tests of Basic Skills math concepts subtest

Group differences significant at the 95% confidence interval.

---

[3] Graphical representations of the data created for this report are based on the data reported in the studies by the authors. All effect sizes in this report were calculated using Cohen's d formula unless otherwise indicated. If standard deviations were not provided by the author of the report, then effect sizes could not be calculated and are not provided. For Cohen's d effect sizes, the results should be interpreted in terms of the following guidelines: 0.2 is equal to a small effect, 0.5 is a moderate effect, and 0.8 and above is a large effect. In practical terms, an effect size of .2 indicates that the average member of the group with the larger mean would score higher than 58% of the members in the group with the lower mean. For an effect size of 0.5, that individual would score higher than 69% of the lower performing group, and for an effect size of 0.8, members in the higher performing group would score higher than 79% of the members in the group with the lower mean score. It should be noted that according to Slavin (1986), a leader in educational research, an effect size of 0.25 is considered educationally significant. Slavin, R.E. (1986). Best-evidence synthesis: An alternative to meta-analysis and traditional reviews. *Educational Researcher, 15*, 5-11.

[4] Neither the number of schools or classrooms is given in the text of the report. However, in a Saxon-created document, it is indicated that there were 56 Saxon classrooms and over 300 Scott Foresman classrooms both spread across K-5 classrooms. This information may be taken from Nguyen's 1994 report, which was a follow-up to the 1993 report. The exact number of students participating in the program and used for the analysis is not known and can only be estimated.

[5] Standard deviation scores were not provided for the student achievement average scores. Therefore, no effect sizes could be calculated.

- These results indicate that, collapsed across grade levels, students who used *Saxon Math* at this Oklahoma school achieved greater gains in their knowledge of math concepts than students using the Scott Foresman product.
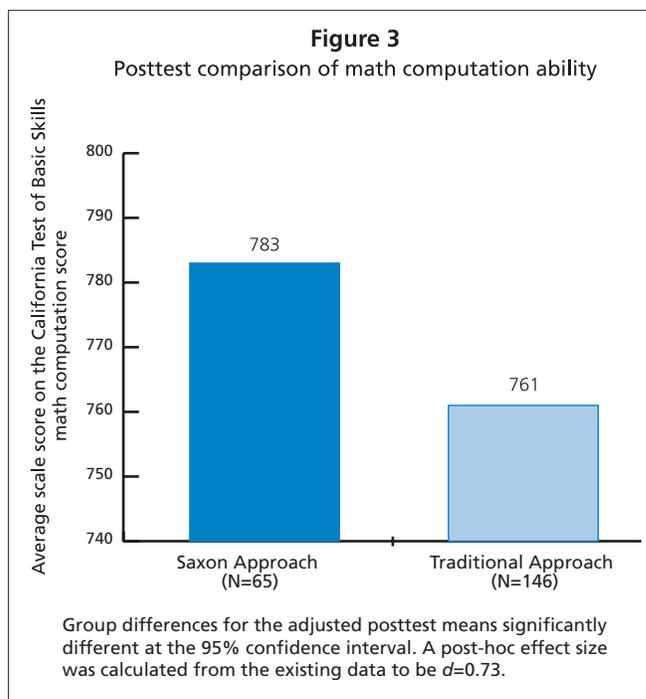
Middle School/Junior High

Sixth Grade – *Math 7/6*

Rentschler, R.V. (1994). *The effects of Saxon's incremental review on computational skills and problem-solving achievement of sixth-grade students*. Dissertation Abstracts International, 56(02), 484A. (UMI No. 9518017).

- This study was conducted as a partial fulfillment to obtain a doctorate degree with the purpose to determine if the incremental approach of *Saxon Math* would help sixth-grade students learn mathematical computations and apply mathematical concepts better than a traditional massed-practice approach.

- Sixth-grade students from two schools in different counties in rural West Virginia were asked to participate in this study based on the schools' choice for math textbook adoption. Sixth-grade classrooms in one school adopted the *Saxon Math 7/6* program (N = 65) while a second school matched in demographics chose to use the Silver Burdett Mathematics textbook (N = 146).

- The California Test of Basic Skills (CTBS), proven to be valid and reliable, was used as the pre- and posttest measure to evaluate mathematics achievement over time. Because of inherent achievement differences between the two schools, the students' previous year CTBS scores were used in the analysis as a covariate in an effort to try to equate the groups' mathematical knowledge. Pretest CTBS scores were examined and were shown to be significantly different for the two groups, further indicating the need to utilize the pretest scores statistically as a covariate.

- Program implantation lasted for one full school year; however, the CTBS was given to the participating students in March of the implementation year after less than a year of implementation. Students' math computation and concept application scores from the CTBS were evaluated in separate analyses.

- As can be seen in Figure 3, the Saxon group's computation posttest mean, adjusted for the pretest differences, was significantly higher than the traditional group's mean. This indicates that the Saxon incremental approach helped to develop students' mathematical computation ability better than the traditional massed-practice approach. However, this pattern was not consistent for the concept application scores. While the Saxon group did score higher on the posttest than the traditional group, the difference was not statistically different.

Figure 3
Posttest comparison of math computation ability

Group differences for the adjusted posttest means significantly different at the 95% confidence interval. A post-hoc effect size was calculated from the existing data to be d=0.73.
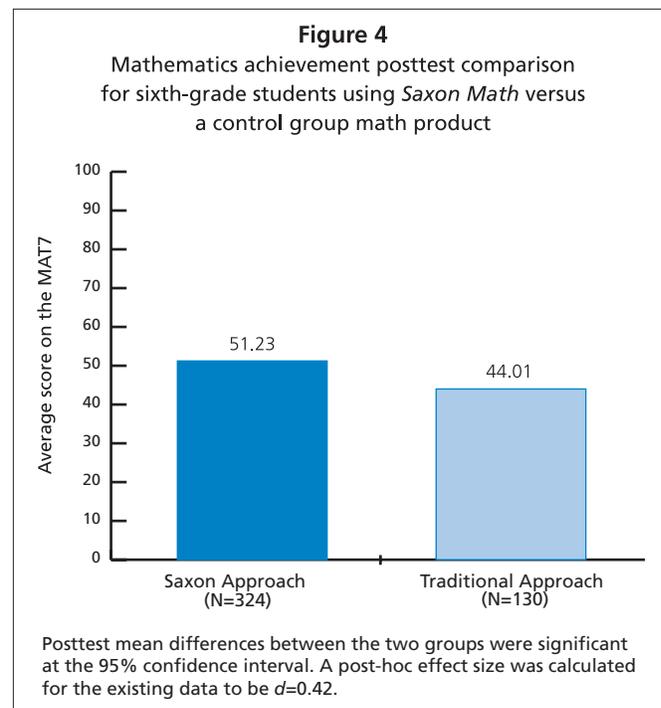
- It was concluded that the *Saxon Math 7/6* program was highly effective at increasing mathematical computation skills, but less effective at increasing application of concept skills. While not specifically noted by the researcher, these findings may have been a result of the timing of the posttest. Because the posttest was given in March and the Saxon program requires an entire year to complete, the full impact of the results may have been underestimated.

Sixth Grade – *Math 6/5*

Lafferty, J.F. (1994). *The links among mathematics text, students' achievement, and students' mathematics anxiety: A comparison of incremental development and traditional texts.* Unpublished Doctoral Dissertation. Widener University, Chester, Pennsylvania.

- This study was conducted as a partial fulfillment for a doctorate degree with the purpose to determine if an incremental method of teaching mathematics leads to better achievement and less anxiety about math performance than a traditional unit-based curriculum.

- Sixth-grade classrooms from two suburban school districts in Pennsylvania were chosen to participate in the study. Individual participation required parental permission. One district comprised of 324 sixth-grade students used the *Saxon Math 6/5* textbook and served as the treatment group. The control group district was smaller, including only 130 sixth-grade students, and used the Addison-Wesley Mathematics textbook.

- The Metropolitan Achievement (MAT) Test 6 (Form M, Intermediate Level) and a revised version of the math anxiety scale from the Fennema-Sherman Mathematics Attitudes Scales were given at the beginning of the year as pretests. Instruction took place daily over the school year in both districts. At the end of the year, the MAT 7 (Form S Intermediate Level 2) and an identical anxiety assessment were given as the posttests. The MAT 7 was administered instead of the MAT 6 because it had been revised during the course of the study. However, the publisher of the MAT 7 indicated that in terms of content, it was highly correlated with the MAT 6.

- Because the initial analysis indicated that the two groups differed significantly on both the pretest scores of the MAT 6 and the anxiety test, an Analysis of Covariance (ANCOVA) was employed to analyze the difference between groups on the posttest scores with the respective pretest scores serving as the covariates.

- Overall, the test scores generally decreased over time for both groups on the MAT. This decrease could be attributed to the change in test version.
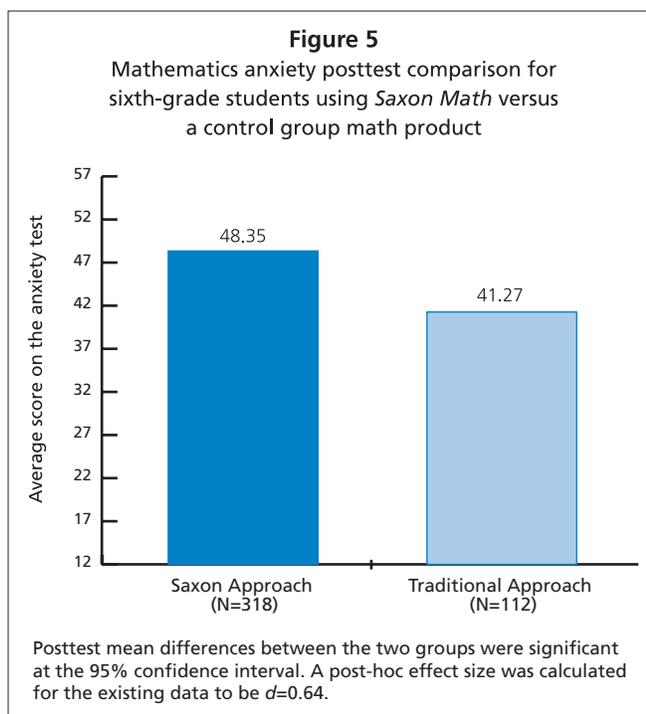
The publisher of the MAT stated that while the versions correlated highly, the MAT 7 was created to be more aligned with NCTM standards and therefore may have contained harder questions or material not covered in either curriculum. Because an ANCOVA analysis was used, only the group difference on the posttest scores was examined after the scores were adjusted for pretest differences. The ANCOVA analysis found that on both the MAT 7 and the anxiety test, the Saxon group significantly outscored the control group (see Figure 4 and Figure 5): thus indicating that the incremental Saxon curriculum facilitated student achievement in mathematics and helped to decrease mathematical anxiety better than the respective unit-driven text.

Figure 4
Mathematics achievement posttest comparison for sixth-grade students using *Saxon Math* versus a control group math product

Posttest mean differences between the two groups were significant at the 95% confidence interval. A post-hoc effect size was calculated for the existing data to be d=0.42.

Sixth Grade – *Math 7/6*

Imrisek, J.P. (1989). *Incremental development: A more effective means of mathematics instruction?* Unpublished master's thesis. Bloomsburg University, Bloomsburg, Pennsylvania.

- This study was conducted as partial fulfillment of a master's degree with the purpose of comparing incremental mathematics instruction to a more traditional unit approach. Two classrooms of sixth-grade students from

Figure 5

Mathematics anxiety posttest comparison for sixth-grade students using *Saxon Math* versus a control group math product

Posttest mean differences between the two groups were significant at the 95% confidence interval. A post-hoc effect size was calculated for the existing data to be *d*=0.64.

- The results from the two assessments were analyzed separately. The two groups did not significantly differ from one another on the SDMT pretest, but by mid-year, after one semester of instruction, the Saxon group had already reached a significantly higher level of mathematical achievement than the control group as measured by the teacher-created end-of-semester exam.

- At the end of the year, the SDMT posttest scores indicated that both groups increased their mathematical ability over time with the Saxon group scoring slightly higher, albeit not significantly higher, than the control group. On the teacher-created final exam, the Saxon group significantly outscored the control group on the overall test, as well as the common questions, and the questions unique to the Saxon curriculum. With regard to the questions unique to the Scott Foresman curriculum, there was no significant difference between the groups. While the control group scored slightly higher on these questions, it was not a significant difference, indicating that the Saxon group performed equally as well on questions from uncovered concepts. See Figure 6 for a graphical representation of the results on the teacher-created final exam.[7]

an urban Pennsylvania city were chosen to participate in the study.[6] One classroom served as the experimental group and used *Saxon Math 7/6*, while the other classroom serving as the control group used Scott Foresman's *Invitation to Mathematics*. The students were randomly assigned to one of the classrooms and both classes were taught by the same instructor eliminating possible teacher effects.

- Using a pretest/posttest control group design, the researcher compared the experimental and control groups on two different assessments. The Stanford Diagnostic Mathematics Test (SDMT) Brown Level, Forms A & B, was used as a general measure of mathematics ability and a teacher-created test, administered as an end-of-semester and final exam, was considered a more direct test of the content material covered by both textbooks.

- The teacher-created semester exam contained only questions that addressed concepts that were common to both curricula. However, the teacher-created final exam, while constructed by the same individuals, differed in its content from the semester exam. The final exam included 45 questions common to both curricula, 17 unique to the Saxon curriculum, and 10 unique to the Scott Foresman curriculum.

- While there was no difference on the SDMT from pre- to posttest between the groups, the researcher concluded that the SDMT was a more general assessment containing easier questions that would be commonly taught in both curricula. Because of this, the researcher believed that the SDMT was not sensitive enough to pick-up the inherent differences between the curricula, whereas the teacher-created exams contained questions that were specifically directed to the content area of the classes, thereby discriminating between the two curricula.

- The overall conclusion was that the incremental approach applied in the Saxon curriculum helped students gain more comprehension

---

[6] N counts for the groups were not specifically provided.

[7] Standard deviation scores were not provided for the student achievement average scores. Therefore, no effect sizes could be calculated.

Figure 6
Comparison of Saxon versus a control group on a mathematics final exam
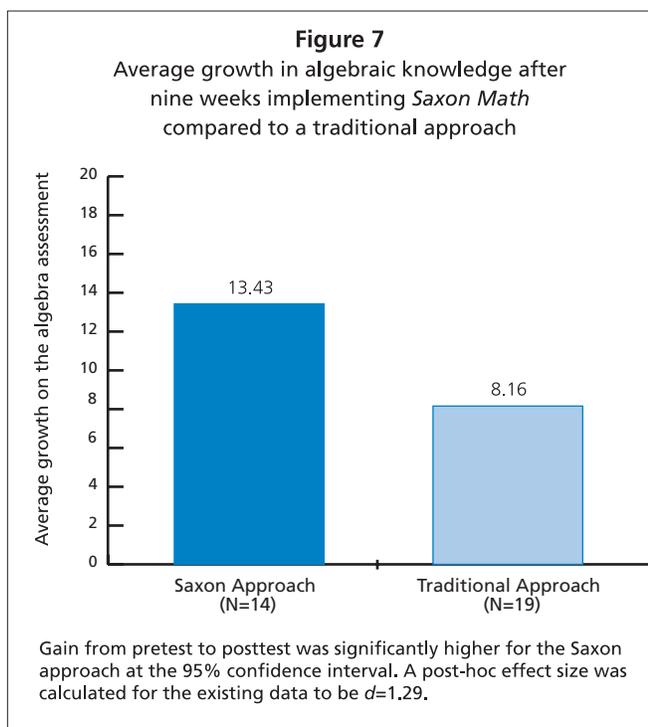divided by type of question
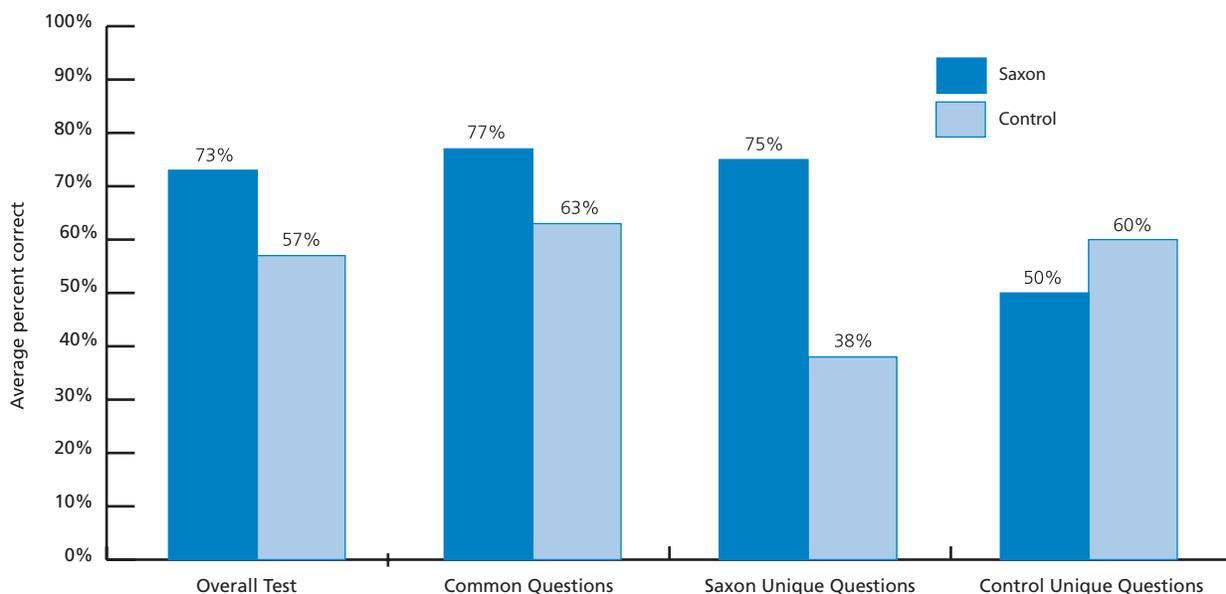
Mean differences were significant between the groups at the 95% confidence interval for all comparisons except for the Control Unique Questions.

and utilization of mathematical concepts than a traditional unit-focused method, which afforded the students the ability to apply this mathematical knowledge and reasoning to untaught concepts.

Eighth Grade – *Algebra I*

Clay, D.W. (1998). *A study to determine the effects of a non-traditional approach to algebra instruction on student achievement.* Unpublished master's thesis, Salem-Teikyo University.

- This study was conducted as partial fulfillment for a master's degree and was designed to compare Saxon with a more traditional approach to increase student achievement in Algebra I.

- Thirty-three eighth-grade students enrolled in an Algebra I class in two schools in rural West Virginia participated in the evaluation. Nineteen of the students served as a control group and used the Fair and Bragg *Algebra* text (published by Prentice Hall). The remaining 14 students were enrolled in a different school and used the *Saxon Math Algebra I* textbook. The two courses were taught by different teachers,

but all students received their respective instruction at the same time of day for the first nine weeks of the school year.

- Locally created criterion-referenced tests designed to measure knowledge of specific algebraic concepts were administered as the pre- and posttest for the evaluation. The tests were specifically designed to cover the concepts taught during the initial sections of both textbooks. The pretest was administered before instruction began and the posttest was given at the end of the nine-week implementation period.

- Results indicated that the Saxon methodology was effective at increasing math achievement. While the Saxon group began the implementation significantly behind the control group, the Saxon approach enabled those students to achieve significantly greater gains from pre- to posttest than the students exposed to the traditional approach (see Figure 7).

- It was concluded that the *Saxon Math Algebra I* program helped students overcome an initial deficiency and bring about greater gains in math achievement than a more traditional approach after only nine weeks of instruction.

Figure 7
Average growth in algebraic knowledge after nine weeks implementing *Saxon Math* compared to a traditional approach

Gain from pretest to posttest was significantly higher for the Saxon approach at the 95% confidence interval. A post-hoc effect size was calculated for the existing data to be $d=1.29$.

Eighth Grade – *Algebra I*

Peters, K.G. (1992). *Skill performance comparability of two algebra programs on an eighth-grade population.* Unpublished doctoral dissertation, University of Nebraska, Lincoln, Nebraska.

- This study was conducted as partial fulfillment for a doctorate degree. The purpose was to compare an incremental method of teaching algebra found in the *Saxon Math Algebra I* program with a traditional massed-practice approach of the National Council of Teachers of Mathematics (NCTM) program formerly known as Chicago Project, which has since become *Everyday Math*.

- Thirty-six math-talented students from a rural high school in Nebraska participated in the study. Math-talented students were those students who had shown a high level of math proficiency in the past as determined by prior scores on the California Achievement Test (CAT), teacher-generated tests, and past class assignments. There were 19 students in the Saxon class and 17 in the NCTM class.

- The school had been using the *Saxon Math* product for three years prior to the study. However, it was interested in evaluating the new NCTM program to determine which program provided students with greater achievement in algebra.

- In addition to math achievement, learner satisfaction, and the effect of prior achievement, knowledge of specific algebraic components were also measured.

- Using the students' prior CAT scores, a district-wide math prognosis test called the Orleans-Hanna Prognostic Test, a learner satisfaction survey, and four researcher-generated criterion unit tests, the achievement of the participating students was compared after a year-long implementation of both programs.

- The pretest scores on the Orleans-Hanna Prognostic Test were compared and no significant difference was found between the groups, indicating that the Saxon group and the NCTM group began the year with comparable math aptitude. After the implementation period, it was found that while both the Saxon and NCTM students raised their overall score, the difference between the groups in terms of gain was not significant. There was equal growth in math aptitude for both programs.

- Prior math achievement as determined by past CAT scores did have an impact on final math achievement at the end of the study; however, there was not a differential effect between groups. For both groups, the higher the previous math score, the more the students achieved throughout the intervention.

- In terms of learner satisfaction, both groups increased their satisfaction equally over time. Neither program provided students with a larger feeling of satisfaction. It was also determined that between learner satisfaction and previous math aptitude, learner satisfaction had a higher impact on final math achievement.

- On the researcher-generated unit tests, the NCTM group performed better on average than the Saxon group on each of the tests. However, no statistical analysis was conducted by the researcher to determine if the differences were significant. Because the data was provided in the original report, a post-hoc analysis of these differences could be conducted, and on three of the four tests, there was no significant

difference between the groups. Only on one of the tests did the NCTM group significantly outperform the Saxon group, indicating that in general, the level of mathematical progress throughout the year was equal in both groups. This indicates that both the *Saxon Math* program and the NCTM program were equally good at contributing to math achievement for this population of math-talented students.

Eighth Grade – *Algebra 1/2*

Crawford, J., & Raia, F. (1986). *Analyses of eighth-grade math texts and achievement.* Oklahoma City, Oklahoma: Oklahoma City Public Schools, Planning, Research, and Evaluation Department.

- This study was designed to evaluate the effectiveness of the *Saxon Math Algebra 1/2* program to increase mathematics achievement for eighth-grade students relative to the Scott Foresman *Mathematics* program.

- The Saxon program was piloted in five middle schools within a school district in Oklahoma City, Oklahoma. Approximately 228[8] eighth-grade students were chosen to use the Saxon program during the 1984-1985 school year. Students' achievement on the CAT from the previous year and the implementation year served as the pre- and posttest scores. The Math Computation, Math Concepts, and Total Math subtests of the CAT were chosen for analysis.

- Student performance was compared in a multitude of ways. Initially, the Saxon users were compared to all other eighth-grade students in the district (approximately 2,447 students)[9]. This comparison found that the students using the Saxon program began the year with higher scores on the CAT, but no significant gains were made by either the Saxon or the comparison group students. Because appropriate statistical analysis could not be conducted due to the nature of the data, it was concluded that based on reviewing the scatterplot of the scores, the Saxon program seemed to benefit the students who had relatively lower pretest scores; however, this assertion was not conclusively determined.

- The second comparison examined the difference in performance between students using the

Saxon program and the students from the same pilot schools using the Scott Foresman text (approximately 697 students). The analysis yielded similar results initially in that the Saxon group had higher pretest scores on the CAT than the Scott Foresman students, and there was no significant gain from pre- to posttest by either group after the year-long implementation. However, in this case, an ANCOVA[10] could be calculated and the analysis found that the resulting adjusted posttest means comparison statistically favored the Saxon program.

- A third comparison, designed to minimize possibly confounding instructor effects, looked only at students with teachers who taught classes both with the Saxon program and the Scott Foresman text (approximately 82 Saxon students and 320 Scott Foresman students). As with the other comparisons, the pretest score significantly favored the Saxon group, but there was no significant difference found in the amount of gain from pre- to posttest between the two groups. However, an ANCOVA resulted in adjusted means that once again favored the Saxon group.

- To further examine these results and equate student pretest differences, the researchers matched equal numbers of students from the Saxon and Scott Foresman groups on their Total Math pretest CAT score. This allowed for comparison only between students who were at similar achievement levels. Because only matching pairs are considered, the N count is equalized but decreased to N = 78 (39 students in each group). Once pretest achievement was matched, posttest differences were found between the groups. For Math Computation, the Saxon group significantly outperformed the Scott Foresman group on the posttest and the amount of gain from pre to posttest. While the posttest
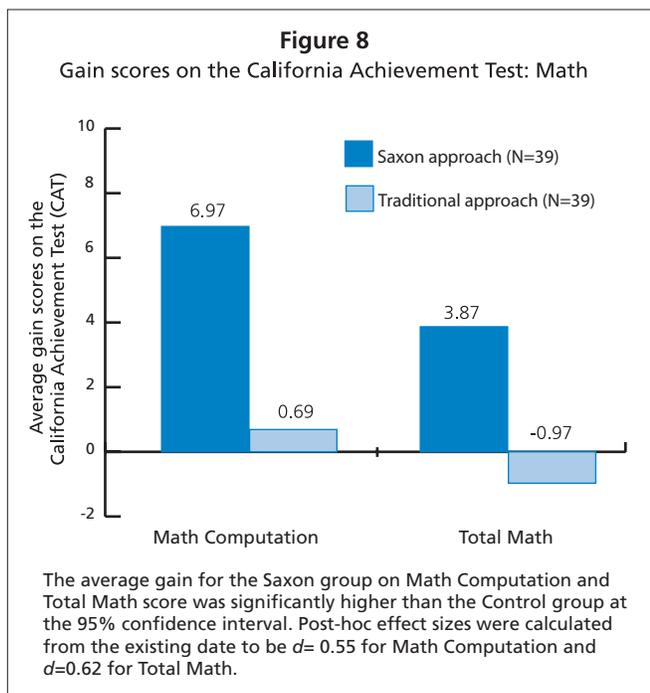
---

[8] The exact number of students was not provided by the author for any of the comparisons.

[9] Not all of the non-Saxon classrooms were using the Scott Foresman *Mathematics* program, but it was the most commonly used program in the district.

[10] The ANCOVA procedure will remove the error that the impact of the differential pretest scores may have on the posttest differences resulting in a more accurate representation of the posttest score differences.

differences were marginally significant for the Math Concepts subtest and favored the Saxon group, the gain from pre- to posttest on the Total Math subtest was statistically significant, with larger gains seen in the Saxon group (see Figure 8 for a graphical representation of the gain score differences on the CAT for the matched groups).

Figure 8
Gain scores on the California Achievement Test: Math

The average gain for the Saxon group on Math Computation and Total Math score was significantly higher than the Control group at the 95% confidence interval. Post-hoc effect sizes were calculated from the existing date to be d= 0.55 for Math Computation and d=0.62 for Total Math.

- As a final analysis, the researchers conducted a content analysis on the CAT and correlated those findings to the concepts covered in both the Saxon program and the Scott Foresman program. The content analysis identified 40 topics on the CAT that were similar to both programs. When the performance of the two groups was examined for CAT items that were covered by both textbooks, the results were significantly different in favor of the Saxon group.
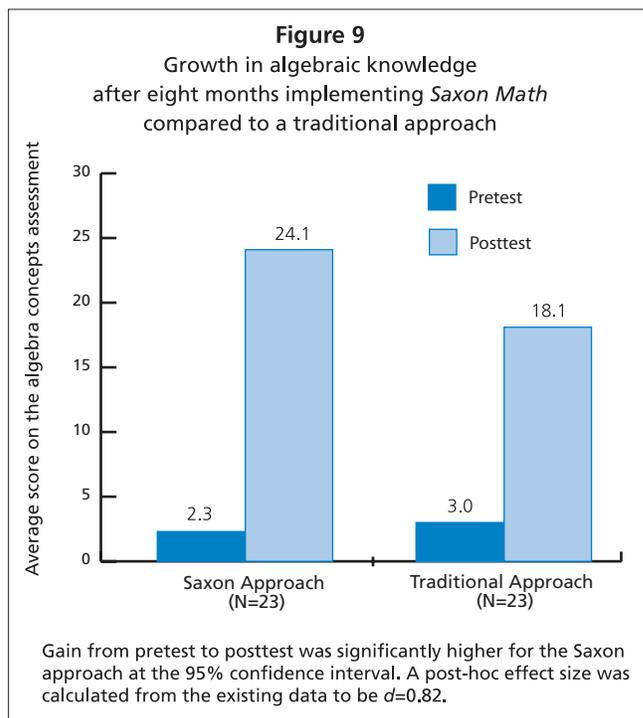
High School

Ninth Grade – *Algebra I*

Williams, D.D. (1986). *The incremental method of teaching Algebra I*. Kansas City: University of Missouri.

- This study reports on an evaluation designed to compare a traditional approach to learning algebraic concepts versus the Saxon incremental method.

- Forty-six ninth-grade students in a rural/ suburban high school in Missouri volunteered for the study based on a desire to take Algebra I. Students were randomly assigned to a classroom using the Saxon methodology or a traditional approach to learning algebra from the Dolciani *Algebra I* textbook. Both groups were taught by the same instructor and each group included 23 students.

- An instructor-created pretest designed to test algebra skills was administered at the beginning of the year. Both groups received daily classroom instruction. Eight months later, the same test was re-administered as a posttest. Students' raw scores on both tests were recorded.

- Average scores on the pretest and posttest for the Saxon group and the traditional group were calculated, as was the amount of gain over the year for each group.

- Statistical analysis revealed that the average gain over the year was significantly higher for the Saxon group than the group using the traditional approach, despite the fact that the Saxon group began the year at a disadvantage. See Figure 9.

Figure 9
Growth in algebraic knowledge after eight months implementing *Saxon Math* compared to a traditional approach

Gain from pretest to posttest was significantly higher for the Saxon approach at the 95% confidence interval. A post-hoc effect size was calculated from the existing data to be d=0.82.

Ninth Grade – *Algebra I*

Mayers, K. (1985). *The effects of using the* Saxon Algebra I *textbook on the achievement of ninth-grade Algebra I students from 1992-1993*. Unpublished doctoral dissertation, Delta State University, Cleveland, Mississippi.
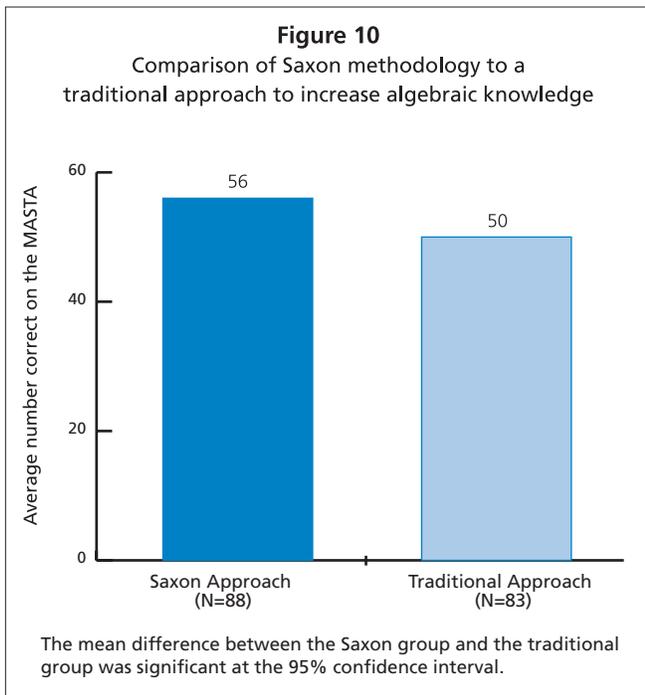
- This study was conducted as partial fulfillment for a doctorate degree with the intended purpose to compare the Saxon incremental method to a traditional unit-based text on a variety of Algebra I skills for ninth-grade students.

- The study compared 171 ninth-grade students enrolled in Algebra I classes at a rural Mississippi high school during a four-year period. Eighty-eight of those students comprised the Saxon group and used the *Saxon Math Algebra I* text, while the remaining 83 used the Dolciani, Brown and Cole *Algebra I* textbook and served as the control group for the study. Group determination was via teacher recommendation, as the original purpose of implementing the Saxon program was to provide a new program for an accelerated group of math learners.

- To control for this inherent group difference, only the comparision group students who had pre-implementation Stanford Achievement Test (SAT) normal curve equivalent (NCE) scores that fell in the same range as those students in the Saxon group were included in the final sample. Other than matching the groups based on SAT NCE scores, no pretest was given to the groups.

- Because the study was conducted over a four-year time period, teacher differences were not controlled. The Saxon group's teachers changed throughout the years while the control group's teacher did not vary.

- The Mississippi Subject Area Test in Algebra I (MSATA), which was created by the Mississippi State Department of Education, was used as the achievement measure for the study. The MSATA was considered a reliable assessment and a valid measure of algebraic concepts.

- In order to determine if the program was successful, students in the Saxon and control groups were compared on 12 variables defined from the categories of problems found on the MSATA:

1. Overall performance on the MSATA.
2. Number of linear equation questions answered correctly.
3. Number of questions correct pertaining to multiplying polynomials.
4. Number of questions correct pertaining to factoring polynomials.
5. Number of questions correct pertaining to simplifying monomials.
6. Number of language of algebra questions answered correctly.
7. Number of questions correct when asked to evaluate or simplify equations.
8. Number of questions correct when asked to plot or identify points on a Cartesian plane.
9. Number of questions correct pertaining to verbal problem translation into equations.
10. Number of questions correct pertaining to identifying opposites, reciprocals, and absolute values.
11. Number of questions correct when asked to add or subtract monomials and polynomials.
12. Number of questions correct when asked to add, subtract, multiply, or divide rational numbers, as well as simplify, multiply, or divide algebraic fractions.

- The analyses indicated that the Saxon group outscored the control group on all variables; however, only the difference between the groups on the overall performance score (see Figure 10) and the scores on categories 8-12 were significant.[11]

- Because each of the variable categories were only represented in some cases by as few as four questions, the researcher considered the student's overall performance on the MSATA as the best measure of algebraic knowledge. It was also noted that the categories on which Saxon significantly outscored the control group were more complex, indicating that as the concepts increased in difficulty, the knowledge base imparted by the Saxon program helped to propel these students to higher achievement levels.

---

[11]Standard deviation scores were not provided for the student achievement average scores; therefore, no effect sizes could be calculated.
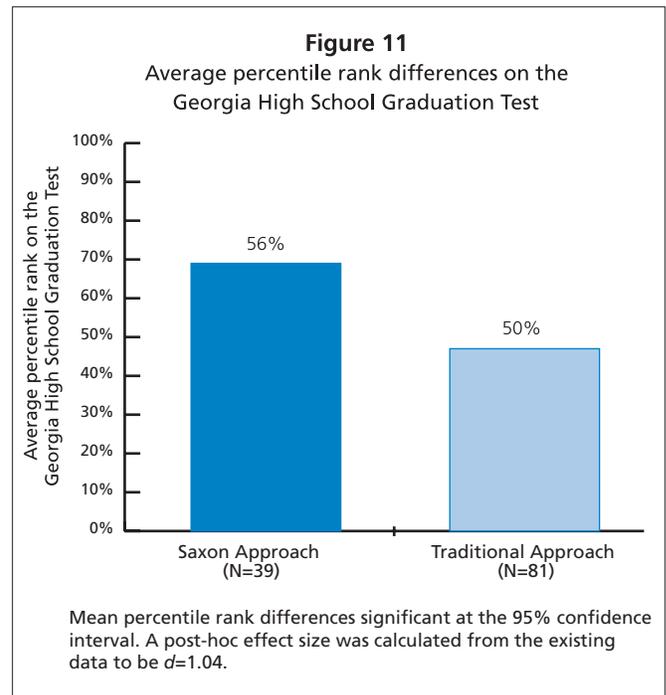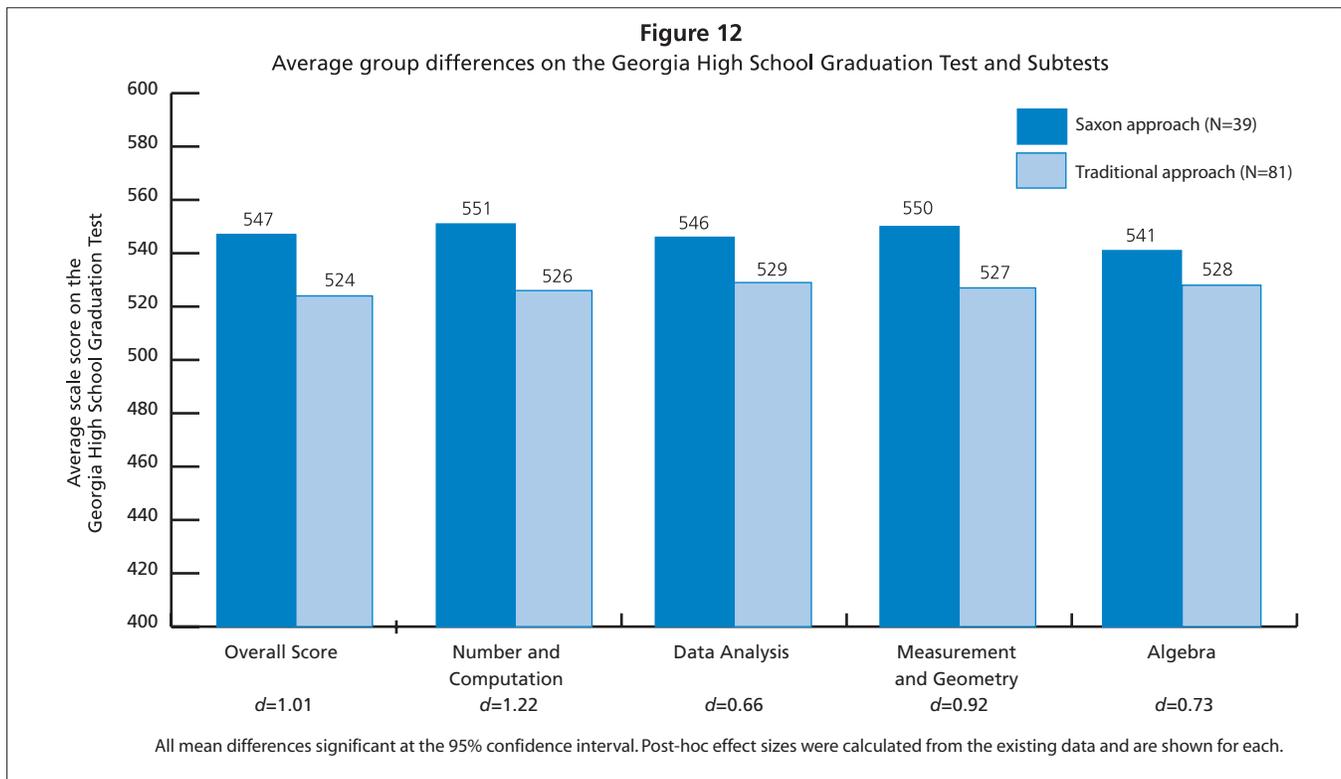
Figure 10
Comparison of Saxon methodology to a
traditional approach to increase algebraic knowledge

The mean difference between the Saxon group and the traditional
group was significant at the 95% confidence interval.

Eleventh Grade – *Algebra 1/2, Algebra I, Algebra II, Advanced Mathematics*

Sanders, B.B. (1997). *The effects of using the Saxon Mathematics method of instruction vs. a traditional method of mathematics instruction on the achievement of high school juniors.* Retrieved September 1, 1999, from www.gsw.ed/~fspaniol/homepage/7420sanders.pdf.

- This study was conducted in an effort to investigate the textbook choices of college preparatory classes at two different high schools in rural Georgia. At one high school, 39 students used textbooks chosen from the Saxon product line, including *Algebra 1/2, Algebra I, Algebra II,* and *Advanced Mathematics*. The other high school included 81 students enrolled in the college preparatory math classes using a variety of different math textbooks that followed a traditional unit-based approach. These included: *Pre-Algebra* (O'Daffer, Clemens, and Charles, 1991), *Algebra I: Structure and Method* (Dolciani, Brown, and Cole, 1988), *Geometry: Applications and Connections* (Burrill, Cummins, Kanold, and Yunker, 1995), and *Integrated Mathematics* (Rubenstein, Craine, and Butts, 1995). The selection of textbook depended upon the curriculum sequence chosen by the student.

- Because the curriculum sequences differed between schools as well as within schools, the study focused the examination at the level of methodological differences between teaching mathematics in a traditional, unit-based approach versus the incremental, spiraling technique of the Saxon product line.

- Given that these preparatory classes were already in place, no formal assignment of students to classes was conducted. Instead, the researcher gained permission to examine data from an assessment already in place for juniors enrolled in these courses.

- Student performance on the Georgia High School Graduation Test, which was given every March at these schools, was used to measure mathematical achievement and to compare the effectiveness of the two types of curricula to influence student achievement in mathematics. Student performance on the assessment was evaluated with an overall score, a percentile ranking, and by subtest. These subtests included: Number and Computation, Data Analysis, Measurement and Geometry, and Algebra.

- Average student scores on the Georgia High School Graduation Test for all six comparisons were calculated for both groups. Figures 11 and 12 detail these findings. The Saxon group scored

Figure 11
Average percentile rank differences on the
Georgia High School Graduation Test

Mean percentile rank differences significant at the 95% confidence interval. A post-hoc effect size was calculated from the existing data to be d=1.04.

Figure 12

Average group differences on the Georgia High School Graduation Test and Subtests

All mean differences significant at the 95% confidence interval. Post-hoc effect sizes were calculated from the existing data and are shown for each.

significantly higher on all comparisons than the group using traditional, unit-based methods to teach mathematics. While there is insufficient data to determine if the students using the Saxon techniques grew at a faster rate over time than the students using the traditional methods, it can be concluded that the students enrolled in the college preparatory classes based on Saxon pedagogy had higher math achievement than students from classes using traditional techniques.
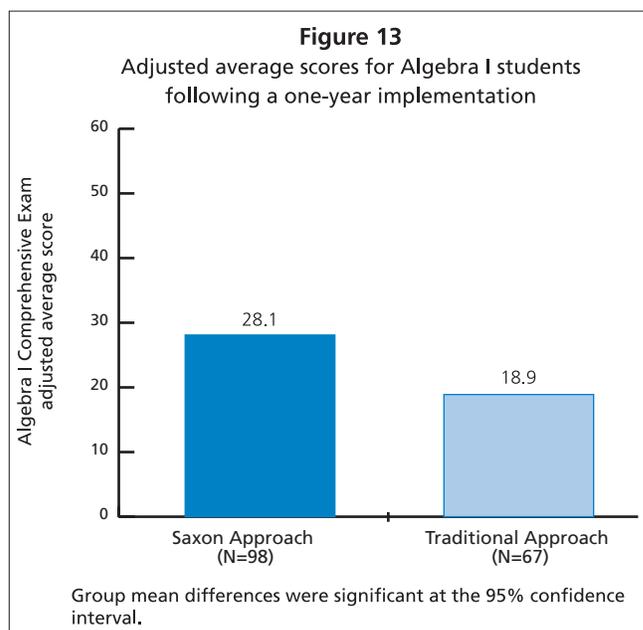
High School Level – *Algebra I*

McBee, M. (1982). *Dolciani versus Saxon: A comparison of two Algebra I textbooks with high school students.* Oklahoma City, OK: Oklahoma City Public Schools.

- This study was the first in the series of studies conducted by the Oklahoma City Schools Planning, Research, and Evaluation Department to investigate the impact of the *Saxon Math* curriculum on student math achievement. For this evaluation, student performance using the *Saxon Algebra I* textbook was compared to the traditional methods presented in the *Dolciani Algebra Structure and Method* book.

- Seven Oklahoma City high schools with two Algebra I classes each participated in the study. One class in each school used the Saxon textbook while the other used the Dolciani textbook and served as the control group. In every case, both classes in the school were taught by the same teacher. Across the schools there were 98 students in the Saxon classrooms, and 67 in the control group. Students were not randomly assigned to classrooms, but textbook assignment to the classrooms was randomly decided.

- After a year-long implementation of the program, a locally created assessment, the Algebra I Comprehensive Exam (ACE), was administered to all students participating in the program. Because the ACE was only given as the posttest following implementation, the students' previous spring California Achievement Test (CAT) scores were used in the analysis to control for possible differences in mathematic achievement prior to the start of the study.

- The ACE was chosen as the posttest measure because it was considered more sensitive to specific algebraic concepts, whereas the CAT is a more generalized test that assesses a wide range of mathematical ability. However, it was

noted that the ACE and CAT had a correlation of 0.76, which is considered a high correlation. The ACE contains questions addressing 21 algebraic concepts that were taught in both programs.

- Because there were inherent student level differences in math achievement at the beginning of the study, the students' CAT scores were used as the covariate in the analysis of the ACE posttest data. ACE scores were compared for the Saxon and control group using their CAT scores to first statistically equate the two groups in ability. The adjusted means on the ACE were 18.9 for the control group and 28.1 for the Saxon group. The difference between these means was highly significant (see Figure 13).[12]

- The ACE scores were additionally broken down and analyzed by algebraic concept. It was found that the Saxon group significantly outscored the control group on 11 of the 21 concepts.[13] Of the remaining comparisons, four were shown to have marginally significant differences between the groups, and for one concept, factoring, the control group significantly outscored the Saxon group. The remaining five comparisons showed no significant difference between the groups.

- The researcher also looked at group achievement difference based on the rate of student absences, the number of missed classes due to assemblies and other school functions, and the amount of completed homework assignments. It was found that the only variable to correlate with student performance was the completion of homework assignments. Within the Saxon group, the correlation between ACE performance and homework completion was $r = 0.28$. Although a relatively small correlation, it nevertheless indicates that for the students using the Saxon program, the more homework assignments that were completed, the higher the achievement scores on the ACE. This relationship did not hold true for the control group.

- Student achievement was also compared in terms of ability grouping. Based on the students' previous year CAT scores, participating students were ranked as low-, medium-, or high-achieving students. The analysis found that the Saxon group scored significantly higher than the control group at each ability grouping. Additionally, there was no differential effect

Figure 13
Adjusted average scores for Algebra I students following a one-year implementation

Group mean differences were significant at the 95% confidence interval.

between the groups, indicating that neither of the algebra programs benefited one ability group over another.

- Because of concerns that were raised about using the CAT to initially equate students on math achievement and then using the ACE as the measure to evaluate student achievement at the end of the study, the researcher examined CAT scores for participating students during the implementation year as well. As with the ACE analysis, students' previous spring CAT scores were used as the covariate to equate the mathematic ability of the two groups. Implementation-year CAT scores were compared between groups, and it was found that the Saxon group scored 63.6, which was significantly higher than the control group's score of 59.5. Although post-hoc, this comparison helps to justify the use of the CAT as the covariate in the original analysis.

- Overall, it was concluded that the *Saxon Algebra I* program was superior to the Dolciani textbook to increase student knowledge of algebraic concepts in Algebra I classrooms.

---

[12]Standard deviation scores were not provided for the student achievement average scores; therefore, no effect sizes could be calculated.

[13]These concepts were: evaluation of expressions, solution of equations with one unknown, rational numbers (fractions), radical expressions, percent word problems, linear equations, simultaneous equations, greatest common factor, addition of like terms, scientific notation, and integral exponents.

Bibliography

Calvery, R., Bell, D., & Wheeler, G. (1993, November). *A comparison of selected second- and third-graders' math achievement: Saxon vs. Holt.* Paper presented at the annual meeting of the Mid-South Educational Research Association, New Orleans, Louisiana.

Clay, D.W. (1998). *A study to determine the effects of a non-traditional approach to algebra instruction on student achievement.* Unpublished master's thesis, Salem-Teikyo University.

Crawford, J., & Raia, F. (1986). *Analyses of eighth-grade math texts and achievement.* Oklahoma City, Oklahoma: Oklahoma City Public Schools, Planning, Research, and Evaluation Department.

Hansen, E., & Greene, K. (2000). *A recipe for math: What's cooking in the classroom?* Retrieved March 3, 2001 from http://www.secondaryenglish.com/recipeformath.html.

Imrisek, J.P. (1989). *Incremental development: A more effective means of mathematics instruction?* Unpublished master's thesis. Bloomsburg University, Bloomsburg, Pennsylvania.

Lafferty, J.F. (1994). *The links among mathematics text, students' achievement, and students' mathematics anxiety: A comparison of incremental development and traditional texts.* Unpublished doctoral dissertation. Widener University, Chester, Pennsylvania.

Mayers, K. (1985). *The effects of using the* Saxon Algebra I *textbook on the achievement of ninth-grade Algebra I students from 1992-1993.* Unpublished doctoral dissertation, Delta State University, Cleveland, Mississippi.

McBee, M. (1982). *Dolciani versus Saxon: A comparison of two Algebra I textbooks with high school students.* Oklahoma City, Oklahoma: Oklahoma City Public Schools.

Nguyen, K., & Elam, P. (1993). *The 1992-1993 Saxon mathematics evaluation report.* Oklahoma City, Oklahoma: Oklahoma City Public Schools.

Nguyen, K., (1994). *The 1993-1994 Saxon mathematics evaluation report.* Oklahoma City, OK: Oklahoma City Public Schools.

Peters, K.G. (1992). *Skill performance comparability of two algebra programs on an eighth-grade population.* Unpublished doctoral dissertation, University of Nebraska, Lincoln, Nebraska.

Rentschler, R.V. (1994). *The effects of Saxon's incremental review on computational skills and problem-solving achievement of sixth-grade students.* Dissertation Abstracts International, 56(02), 484A. (UMI No. 9518017).

Sanders, B.B. (1997). *The effects of using the Saxon Mathematics method of instruction vs. a traditional method of mathematics instruction on the achievement of high school juniors.* Retrieved September 1, 1999, from www.gsw.ed/~fspaniol/homepage/7420sanders.pdf.

Williams, D.D. (1986). *The incremental method of teaching Algebra I.* Kansas City: University of Missouri.

Harcourt Achieve

Rigby • Saxon • Steck-Vaughn

www.HarcourtAchieve.com
1.800.531.5015