Using the FCAT Explorer to Improve 5th Graders' Math FCAT Performance

by

Bahia Naime-Diefenbach, Ph.D.
(Bahia@fdt.net)

and

Christopher B. Sullivan, Ph.D.
(Christopher@imageresearch.com)

Paper presented at the Florida Educational Research Association Annual Conference
November 19 - 21, 2003, Orlando, Florida

Using the FCAT Explorer to Improve 5th Graders' Math FCAT Performance

Introduction

Computer-based educational materials are routinely used to prepare students for standardized tests such as the Florida Comprehensive Assessment Test (FCAT). Educators, who need their investment in software to work, keenly note the efficacy of each educational software program. This study is the result of a practical question about the efficacy of the Department of Education's FCAT Explorer in helping prepare students for the Florida Comprehensive Assessment Test. The program is intentionally constructed based on a model of learning that combines practice and feedback to instruct students in the math benchmarks. The practical research question is "does using the practice software actually help a student to perform well on the FCAT?"

This study examines one software program, the Florida Department of Education's FCAT Explorer: 5th Grade Math (5GM), which is an interactive program available to all Florida public schools over the Internet. The FCAT Explorer 5GM is a math program that enables students to practice the benchmarks and skills for the 5th grade math FCAT. The study demonstrates that the FCAT Explorer 5GM's incremental learning steps, learning guidance feedback and the extra practice of trying the math item again, enhance students' acquiring math strategies to solve math problems, and this learning is carried over to the students' performance on the FCAT.

The analysis reported here was undertaken with research staff in the Leon County Schools district office. The research was conducted at no charge to the school district by Image Research. Its intent is to compare student performance on the FCAT Explorer 5GM program during the 2001-2002 school year with each student's subsequent 2002 fifth grade math FCAT score. The paper investigates the effects of using the FCAT Explorer 5GM to prepare students for the Florida Comprehensive Assessment Test (FCAT). We argue that the main effect of students' use of the FCAT Explorer 5GM is to enhance their learning of the math benchmarks and their mastery of 5th grade math skills. The final outcome of using the FCAT Explorer is seen in students' performance on the FCAT itself.

Learning Mathematics in a Multimedia Environment

Currently, with so many advanced technologies and computers available to them, students feel comfortable using these media for entertainment and learning. Technology classroom settings become captivating and intriguing learning environments. Although multimedia-teaching technologies still come with a high cost (Fahy, 2000), multimedia materials show promising effects on students' acquisition of knowledge. Multimedia programs can enhance teaching and learning for today's diverse students (Torrez, 2000) with a variety of learning styles (Hawkins, 1993; Herrington & Oliver, 1999; Schank, 1993), creating greater learning environments and ameliorating learning outcomes (Goldenberg & Cuoco, 1996, Schifter, 1997; Russell, 1997). Through the years, these innovative multimedia learning environments and technology-enriched classrooms have had a positive effect on students' acquisition of higher-order thinking skills (Hopson, Simms, & Knezek, 2001-2002; Paolucci, 1998; Schank, 1993). In addition, some software can support complex learning in math as well as teach problem solving to students who struggle with learning difficulties (Nicaise, 1997; Babbitt and Miller, 1996).

An important feature of multimedia software applications is their interactivity with their users and their ability to provide important feedback, with a varying amount of help and support (Hannafin & Scott, 1998). Therefore, when selecting the appropriate software, teachers and administrators need to carefully evaluate whether the software is consistent with the expected curriculum and learning outcomes (Goyne, McDonough, Sharon, Padgett, Dara, 2000).

Learning Mathematics using Real-Life Situations

Learning mathematics requires students to make sense of abstract mathematical concepts. Gaining knowledge of mathematics means learning to use and understand its language, its symbols and its procedures. Many learners experience barriers in visualizing mathematical concepts. Students must learn how to construct knowledge, conceptualize problems, and develop problem-solving skills. By applying mathematics skills and theories in real life situations, students establish connections between school learning and their interests outside school. This process enhances critical thinking and mathematics reasoning and learning skills and improves the retention and transfer of learning (Schoenfeld, 1992; Verzoni, 1997). Presenting new math entities to the students in a variety of meaningful and authentic ways, and introducing relevant relations, might permit an easier pathway to learning than just teaching for the construction of

cognitive units. This might change dramatically the overall attitude of students toward mathematics. Simulations of real-life situations enable the learners to acquire knowledge about mathematics skills by way of inferences from empirical knowledge. Simulations also introduce concepts based on their experiences with cases from everyday life. More importantly, learners would be able to carry those fundamentals to unfamiliar situations.

Immediate Learning Feedback

Multimedia learning environments enhance the learning of mathematical skills by providing the learner with interactivity, immediate feedback, control of the pace of instruction, and individualized learning (Hawkins, 1993; Naime-Diefenbach & Sullivan, 2001). While feedback seems to be important in the enhancement of learning, research indicates that this is true only under certain conditions (Cooper, 1998; Khine, 1996; Kim and Sharp, 2000). For example, educators should establish a supportive environment in which students are comfortable giving a wrong answer, and more importantly, the nature and not the quantity of the provided feedback is critical (Golding, 2003). In a learning situation, feedback is the information obtained by students regarding the accuracy of their performance in a learning task. Feedback can be categorized according to its functions and characteristics (Dempsey and Sales, 1993). These include two main types: 1) knowledge of the results, which provides responses such as “right” or “wrong,” without giving the correct answer; and 2) elaborative feedback, which contains the result of a learner’s response, the reason for “why” the response was wrong and the correct answer.

Studies have examined immediacy of feedback, and the amount of information presented in the feedback. Researchers agree that elaborative and informative feedback does benefit learning and enhances performance for several types of learning tasks, specifically when the feedback relates to the correct answer (Kulhavy and Wager, 1993; Khine, 1996; Keller, 1987; Naime-Diefenbach & Sullivan, 2001).

The FCAT Explorer: 5th Grade Math Program

The FCAT Explorer programs comprise a suite of three math and five reading programs funded and endorsed by the Florida Department of Education for FCAT practice (www.fcatexplorer.com). The FCAT Explorer is provided via the Internet to all public school students in Florida at no charge. The programs can be used at school or at home, though students

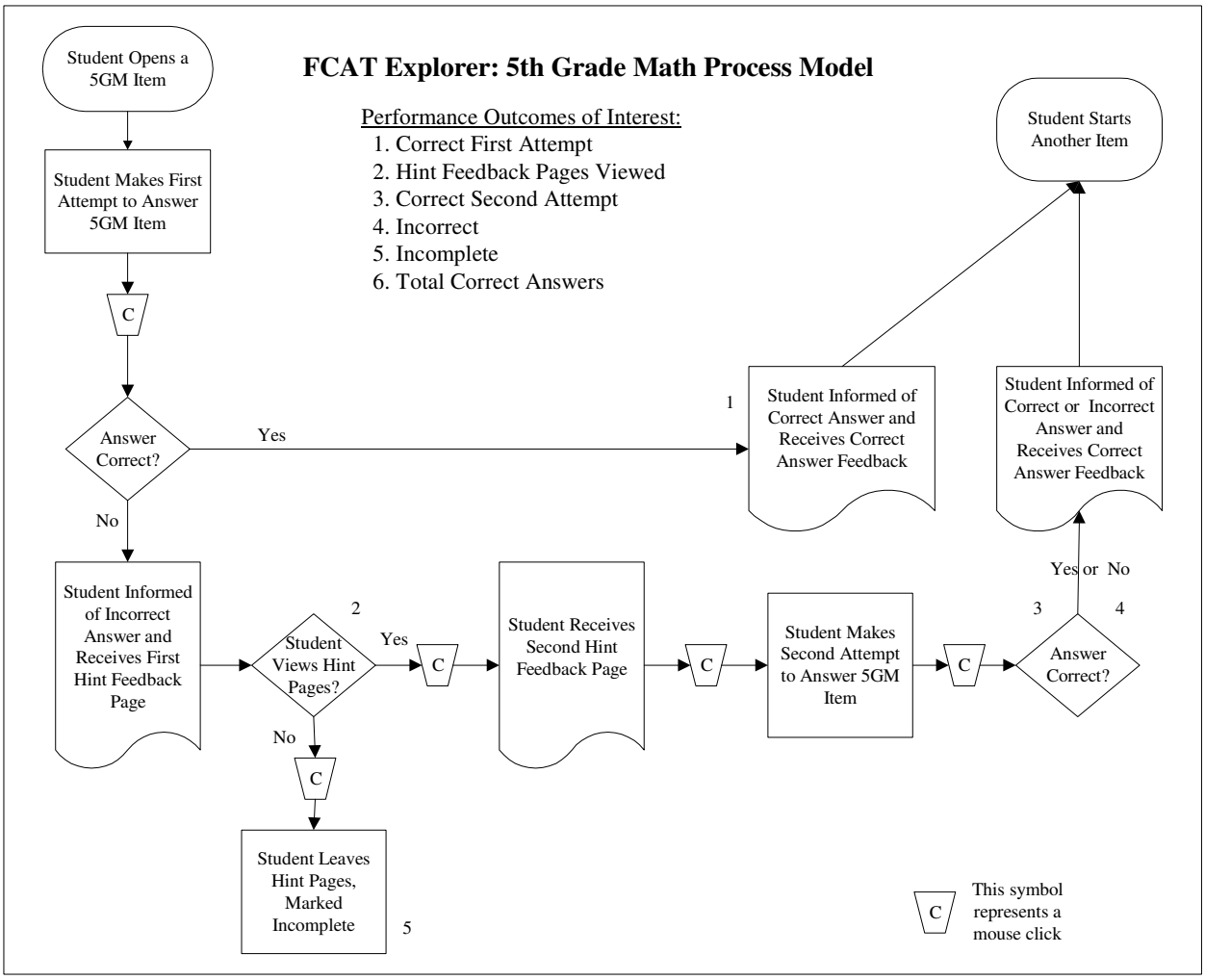
need a username and password to access the program. The program under consideration in this study is the FCAT Explorer: 5th Grade Math (5GM) program. The FCAT Explorer 5GM program consists of a series of 148 math items in the form of math word problems written to the benchmarks at the 5th grade math level, as specified in the Sunshine State Standards. Students who work through all of the items in FCAT Explorer 5GM will practice items related to all of the 5th grade math benchmarks tested in the FCAT.

The FCAT Explorer 5GM was created using the advice of expert math teachers, education specialists, instructional designers, and testing professionals. The educational materials were designed around current cognitive learning theories and strategies, and motivational models (Bloom, 1956; Dick and Carey, 1990; Gagne, 1987; Keller, 1987, Naime-Diefenbach, 1991). The instructional materials in the FCAT Explorer 5GM focus on reinforcing the strategies and knowledge needed to enhance learning of 5th grade math benchmarks and skills. One approach of the FCAT Explorer 5GM is the use of real-life applications of math concepts in each word problem.

The items in the FCAT Explorer are formatted as multiple-choice or gridded response problems, with a few charting problems included. Each item consists of a stem; an illustration visually depicting the word problem; and four distracters for multiple-choice items or a grid for gridded items. Each student is given two attempts to answer the item correctly, and is then brought to the next item in the sequence. After the first attempt to answer the item, elaborative feedback is provided to help the student understand and solve the problem correctly. Students also receive a correct answer explanation after they answer the item, whether correctly or incorrectly. The combination of practice and feedback in this heuristic process is expected to help the student learn math concepts related to specific benchmarks and enhance mathematical skills and strategies, thus, he or she will perform better on the math FCAT.

The progress of a student through the different practice and feedback steps in the FCAT Explorer 5GM is illustrated in Figure 1. A student initiates the answer cycle by reading a math item and attempting an answer; the cycle ends with the student receiving correct answer feedback, then a new item. The answer cycle follows the same logic path for all items.

Figure 1. FCAT Explorer Process Model



The student must initially work through a set of menus in the FCAT Explorer to bring up an item for practice. After reading the math problem the student enters the answer, either by selecting a letter (multiple choice) or typing in the correct number (gridded response). The program assesses the response and depending upon whether the student's answer is right or wrong, presents different types of elaborative feedbacks.

When students answer the item correctly, they receive a "correct answer explanation" page from the FCAT Explorer. This page consists of an animated marquee that spells out "CORRECT" and the answer to the problem, reinforcing one or more solution techniques and appropriate math strategies. The student can then continue to the next item, or back to the main

menu. Correct answers on first attempt are counted as a performance outcome, as indicated by the number 1 in the figure. We assume that students who answer the item correctly know the material covered in the item.

If the student answers the item incorrectly the FCAT Explorer responds to the answer selected with "common error feedback" that addresses the particular common mistake associated with the selected distracter. The feedback comes to the student in two pages. The first page explains how to read the question and the second page provides a hint that directs the student's thinking toward reasoning through the problem. The number of hint feedback pages presented is counted as a performance outcome, as indicated by the number 2 in the figure. The count of hint feedback pages provides a measure of the learning opportunities made available to the student. We assume that the students read the hint feedback pages, before attempting a second answer.

After reading the feedback pages, the student is allowed to attempt the item a second time. For multiple choice items, the first answer is grayed out and the student's chance of guessing correctly improves to one out of three. If the student answers the item correctly, he or she receives the animated marquee and the "correct answer explanation" page. The number of times a student answers correctly on second attempt is counted as a performance outcome and is indicated by the number 3 in the figure. We assume that answering correctly on second attempt indicates that the student has read and acquired the math skills presented in the hint feedback pages to get the right answer.

If a student answers the item incorrectly on second attempt, the "correct answer explanation" page comes up, with a marquee that states the answer is incorrect. Otherwise, there is no difference between the two pages. Incorrect answers are counted as a performance outcome, as shown at number 4 in the figure. We assume that the student was not able to learn enough from the hint feedback pages to help him or her answer the question.

A student also has the option of not attempting to answer the item a second time. The decision to opt out of the item occurs within the hint feedback pages, generally by going back to the menu instead of going forward to a second attempt at the item. The number of incomplete answers is counted as a performance outcome, as shown by the number 5 in the figure.

The pattern of using "common error feedback" and "correct answer explanation" continues consistently through the entire program. This sequence of practice answers followed by guidance

feedback presents an opportunity for students to learn. The expected total outcome is increased knowledge of the math benchmarks and better performance on the math FCAT.

Methodology for Analyzing FCAT Explorer Effects

The interest behind this study was to determine the efficacy of using the FCAT Explorer 5GM to practice for the math FCAT. Being able to compare the FCAT scores of Leon County students with their performance on the FCAT Explorer was a great opportunity to test the research question: "does using the FCAT Explorer help a student to perform well on the FCAT?" If the instructional theory and design that went into building the FCAT Explorer 5GM is correct, there should be a positive relationship between using the FCAT Explorer 5GM and math FCAT performance. This study sets out to determine that relationship, using the performance outcome indicators available from the FCAT Explorer database, plus the school information and FCAT scores of students in the sample.

Sample Population

The sample of students in this study was drawn from the population of Leon County students who used the FCAT Explorer 5GM program during the 2001-2002 school year. The students were included in the study if they had used FCAT Explorer 5GM to answer at least one item in the program (the average student in the study answered 41 items).

The FCAT Explorer 5GM activity records for each Leon County student were extracted directly from the FCAT Explorer database for the period between September 1, 2001 and March 15, 2002. The records for each student were then summarized into performance totals to provide information on how many items a student attempted, how many were answered correctly on first try and so on. This process yielded a database of 1,824 students with summary records of their FCAT Explorer activity.

The student database was passed on to the Leon County School District research office where the records were joined to the Leon County student database to provide information on grade level and curriculum group, math FCAT scores for 2001 and 2002, achievement levels of students for 2001 and 2002 and other assessment data. The resulting dataset then had all student identification removed to make the dataset anonymous prior to analysis.

Of the 1,824 students, 801 were in grades other than 5th grade, so these students were dropped from the sample. Next, to reduce differences in the sample, we decided to drop 107 students who were classified as exceptional students (ESE) or English as second language students (ESL), leaving a final sample of 916 students. The sample was spread across 20 elementary schools in the county, and included schools with 2001 school grades of A, B, C and N (No Grade). Schools in Leon County that did not use the FCAT Explorer 5GM are not included in this sample. Table 1 presents a breakdown of the schools and students in the sample population by school grade. It also compares average activity on the FCAT Explorer and the average math FCAT scores for 2001 and 2002 by school grade.

Table 1. Sample Population, FCAT Explorer Activity and FCAT Scores by School Grade

2001 School Grade	2001-2002 Number of Schools in Sample	2001-2002 Number of Students in Sample	2001-02 Average Number of 5GM Items Attempted	2001 Average Math FCAT Score of Sample	2002 Average Math FCAT Score of Sample
Grade A	8	458	50.3	341.7	356.4
Grade B	5	227	28.5	325.7	347.7
Grade C	4	168	38.2	278.0	306.9
Grade N	3	63	24.0	300.2	315.7
Total	20	916	40.8	323.3	342.4

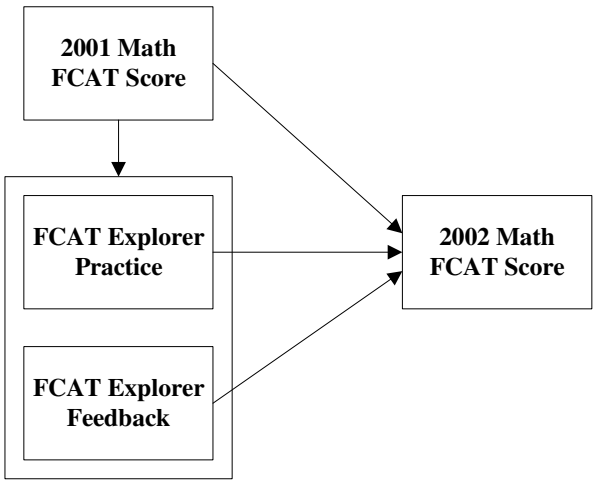
More than half of the students in the FCAT Explorer 5GM sample attend A schools and the number decreases with grade level. This disparity could mean that students in A schools had more access to, or were more inclined to use the program; or it could reflect the relatively large number of A schools in Leon County (35% of all elementary schools in 2001). Program usage is highest for students in A schools, drops by almost half for B schools then increases sharply for C schools. This pattern is consistent for FCAT Explorer 5GM usage across the state in 2001-2002. Finally, the FCAT scores decrease with school grade level, as would be expected.

A constraint of this sample is that there is no comparison control sample of students who did not use the FCAT Explorer, so there is no way to compare changes in FCAT scores between 2001 and 2002 for those students. This limitation does not take away from the value of this FCAT Explorer sample, but it does prevent us from making some important comparisons between students who did use the FCAT Explorer and those who did not.

Indicators of Interest in the Effects Model

The expectation of the FCAT Explorer effects model is simply put: "Using the FCAT Explorer has a significant, positive effect on subsequent FCAT performance." Students answering items in the FCAT Explorer 5GM receive both practice in working through an item and feedback in the correct answer explanation and hint pages. An effects model should account for these two different activities and should predict that both practice and feedback will have a positive effect on learning and on subsequent FCAT performance. The FCAT Explorer effects model is pictured in Figure 2, which shows major influences on a the 2002 math FCAT score of a sample student: his or her math score in 2001 and activity on the FCAT Explorer 5GM. Performance on the 2001 FCAT is related to a student's activity on the FCAT Explorer 5GM, as discussed below, and should predict that student's performance in 2002. The effect of the 5GM activity on a student's knowledge and math skill should also provide an additional explanatory factor for FCAT performance. This study examines the effect of FCAT Explorer activity on 2002 FCAT performance, then adds the predictor of the 2001 FCAT score to compare just how much explanatory power the use of the 5GM adds to the effects model.

Figure 2. FCAT Explorer Effects Model



One constraint of the FCAT Explorer data is that it only counts of the number of items attempted and the outcome of those attempts, as indicated in the FCAT Explorer Process Model. There are no direct measures of student attention to the feedback pages, or of learning across all the practice sessions. We use the construct of correct answers on first attempt as an indicator of the sufficiency of a student's knowledge and skill level. We use the correct answer on 2nd

attempt only as an indicator of immediate learning from the guidance feedback, and the incorrect answer as an indicator of too little knowledge and skill despite the guidance feedback. The measure for long-term learning is the 2002 math FCAT score. We assume that the classroom side of teaching is similar for all students and we control for the 2001 FCAT scores by comparing students according to the 2001 school grade and by the achievement level in which they were placed from their 2001 FCAT scores.

The raw score indicators for FCAT Explorer Practice and Feedback constructs are made up of the count of outcomes from attempts to answer an item, as shown in Figure 1. The variables used to test the effects model are percentage indicators derived from these raw scores, that allow a comparison of performance based on the students' own level of activity. The Practice and Feedback indicators, including both raw score variables and the constructs based on percentage calculations, and are shown in Table 2 and Table 3 below.

Table 2. FCAT Explorer Practice Indicators

Raw Score Indicators of Practice	**Practice Constructs Based on Raw Scores**
Total Items Attempted	N/A
Number of Correct Answers	Percent of Correct Answers (Total Correct / Total Answers)
Number of Correct Answers on 1st Attempt	Percent of Correct Answers on 1st Attempt (Correct 1st / Total Answers)
Number of Correct Answers on 2nd Attempt All Answers	Percent of Correct Answers on 2nd Attempt of All Items (Correct 2nd / Total Answers)

The indicators for FCAT Explorer Feedback are indirect, since feedback pages are not counted directly by the FCAT Explorer. The feedback measures are derived from the practice attempts, and represent proxies of the measures. Each student receives two hint feedback pages after answering an item incorrectly; all students receive a correct answer feedback page when they finish their attempt. Receiving feedback pages does not ensure that the student learns what he or she needs to know, so we use the percent of correct answers after hint feedback to represent the proportion of hint feedback pages that successfully imparted the math information needed by the student. These indicators include:

Table 3. FCAT Explorer Feedback Indicators

Raw Score Indicators of Feedback	Feedback Constructs Based on Raw Scores
Number of Hint Feedback Pages Presented to Student	N/A
Number of Correct Answer Feedback Pages Presented to Student	N/A
Number of Correct Answers on 2nd Attempt Only	Percent of Correct Answers on 2nd Attempt Only (Correct 2nd / Total 2nd Attempt)
Number of Incorrect Answers	Percent of Incorrect Answers (Incorrect / Total Answers)
Number of Incomplete Answers	Percent of Incomplete Attempts (Incomplete / Total Answers)

The percentages and counts of the practice indicators and feedback variables are put into the effects model as separate influences for student learning. All of these variables are used to explain the student's performance on the 2002 math FCAT, though no one variable is expected to explain all of the variance in the FCAT scores.

Research Hypotheses:

The research model proposes direct, positive effects from practicing items in the FCAT Explorer 5GM and from reading the feedback pages. From the effects model, we would expect to confirm each of the following research hypotheses below. The total number of practice items attempted is a measure of how many times a student runs through the FCAT Explorer 5GM practice/feedback cycle. The more items a student works through, the better he or she should perform on the FCAT. Hence:

1. The number of practice items that a student attempts will be positively related to the student's FCAT score.

The best measure of whether a student understands math benchmarks is if he or she can answer a math problem correctly, either on the first or second attempt. The 1st correct answer indicates knowledge of the subject; the 2nd correct answer indicates that learning occurred from

the hint feedback pages. Students with a greater proportion of correct answers should perform well on the FCAT. Hence:

2. The percentage of correct answers that a student makes will be positively related to the student's FCAT score.

Conversely, if a student answers an item incorrectly then the student did not learn enough from the hint feedback pages to solve the math problem. Students with a high proportion of incorrect answers should not perform well on the FCAT. Hence:

3. The percentage of incorrect answers that a student makes will be negatively related to the student's FCAT score.

Finally, working through the feedback pages should help a student learn about math concepts. If the feedback pages function correctly, then the more feedback opportunities a student has to learn, the better that student should perform on the FCAT. Hence:

4. The number of feedback pages that a student views will be positively related to the student's FCAT score.

Testing the FCAT Explorer Effects Model

The test of the FCAT Explorer effects model should verify each of the research hypotheses listed above, and should clearly indicate that both practice and feedback in the FCAT Explorer 5GM program explain performance success in the math FCAT. Further, the test of the model should indicate whether the grade of the school or the achievement level of the student plays a role in the success of the FCAT Explorer 5GM. Testing the FCAT Explorer effects model is carried out in several steps, looking initially at the relationship between the 5GM components and the 2002 FCAT score, then including the external factor of the 2001 FCAT scores.

The model is first examined by separating out students by school grade and by achievement level to control for different skill levels. The distinctions among skill level might be critical to understanding how the FCAT Explorer 5GM, and software programs like it, can influence children's math learning and performance. For example, Table 1 shows that students in 2001 A schools answer more FCAT Explorer 5GM items than students in any other level; students in the highest achievement level also get more items correct than students in other achievement levels. The analysis then presents Pearson correlations to estimate the degree of relationship between the practice and feedback variables in the effects model and the 2002 math FCAT scores. Those

variables that have little explanatory power in the model are eliminated and the effects model variables that demonstrate a significant correlation with the 2002 math FCAT scores are entered into a multiple regression model to determine the explained variance in the FCAT scores attributable to 5GM practice and feedback variables. Finally, the FCAT Explorer 5GM components are compared to the 2001 math FCAT scores to confirm their explanatory power within the larger effects model shown in Figure 2.

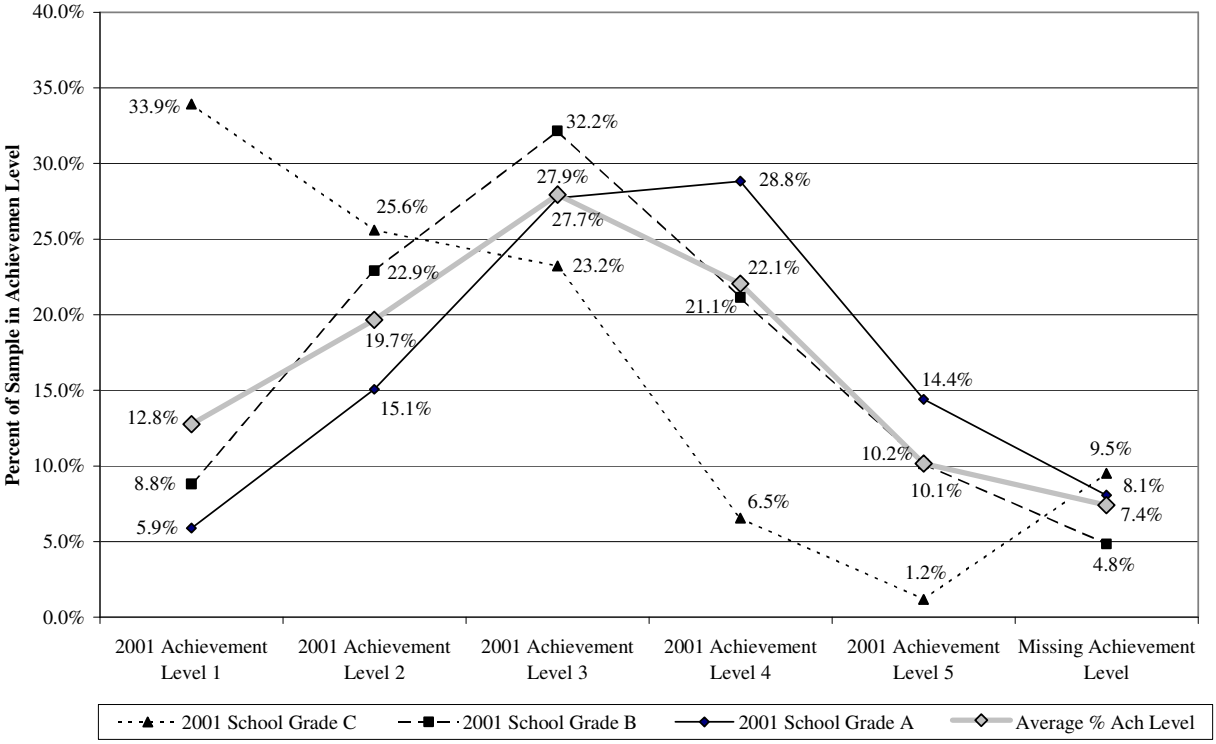
Results of the FCAT Explorer 5GM Effects Model Analysis

The analysis of Leon County students sets out to measure the effects of the practice and feedback components contained in the FCAT Explorer 5GM program on the performance of students who take the math FCAT. The population sample of students who used the program contains students in all achievement levels, across all school grades. We assume that students' performance on the FCAT Explorer will differ according to both school grade and achievement level, so we begin the analysis with summary statistics tabulating the different levels of practice and feedback for students in each school grade and achievement level.

Comparison of FCAT Explorer 5GM Activity with Achievement Level by School Grade

The distribution of the sample students across achievement levels approximates a normal curve, as shown in gray in Figure 2, with about 30% of all students in Achievement Level 3, 20% in Achievement Levels 2 and 4, and about 10% in Achievement Levels 1 and 5. When students are broken out by school grade, however, the normal curve breaks up into three different curves, also shown in Figure 3. Here, the percent of students in each achievement level is compared by school grade. Complete records of these data are given in Appendix A. In this chart it is clear that students in the C schools fall into the two lowest achievement levels to a much greater degree (59.5%) than students in either the B (31.7%) or A schools (21.0%). The C schools have fewer students in the top two achievement levels and the B schools remain consistently below the A schools in the percentage of students in the top two levels. These differences in the student sample suggest that student performance on the FCAT Explorer will also differ across school grades.

Figure 3. Percentage Distribution of Students in Achievement Levels by School Grade



A comparison table providing details of how students performed on each of the practice and feedback variables, by school grade and achievement group, is provided in Appendix B. This analysis focuses on just two comparisons to demonstrate sample differences: the percent correct on first attempt, and the percent correct after the hint feedback. Figure 4 compares the percentage of times students answered correctly on first attempt in each achievement level, and by school grade. Student performance at each achievement level is remarkably similar, across all school grades, with the lowest achievement levels distinctly lower than the top achievement level.

As expected, the A school students perform better than students in the other school grades. The linearity of the graph suggests that the students' performance on first attempt to answer a math item parallels their performance on the 2001 math FCAT; students with better FCAT scores in 2001 are able to answer more questions correctly the first time than students who did poorly. The results of answering the math problem after receiving hint feedback is also apparently related to achievement level, as Figure 5 shows with the percent of times students answered the item correctly on the second attempt.

Figure 4. Comparison of Correct Answers on 1st Attempt (P-Value of 1st Attempt) by Achievement Level and School Grade

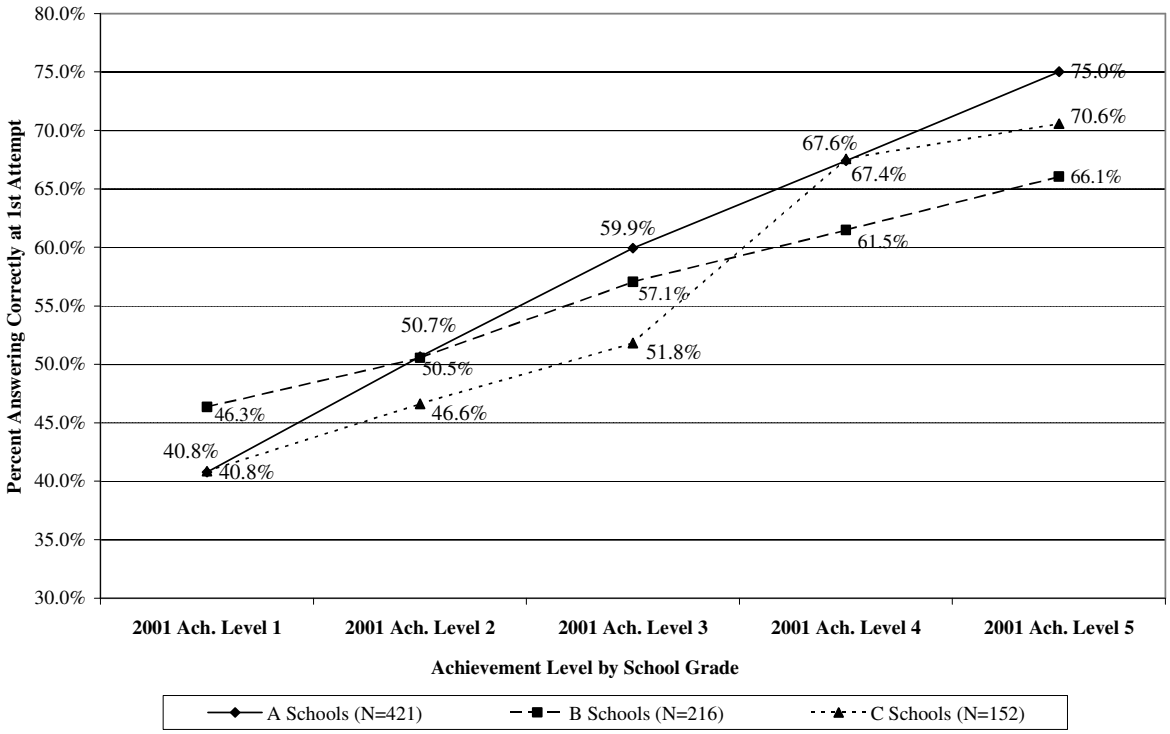
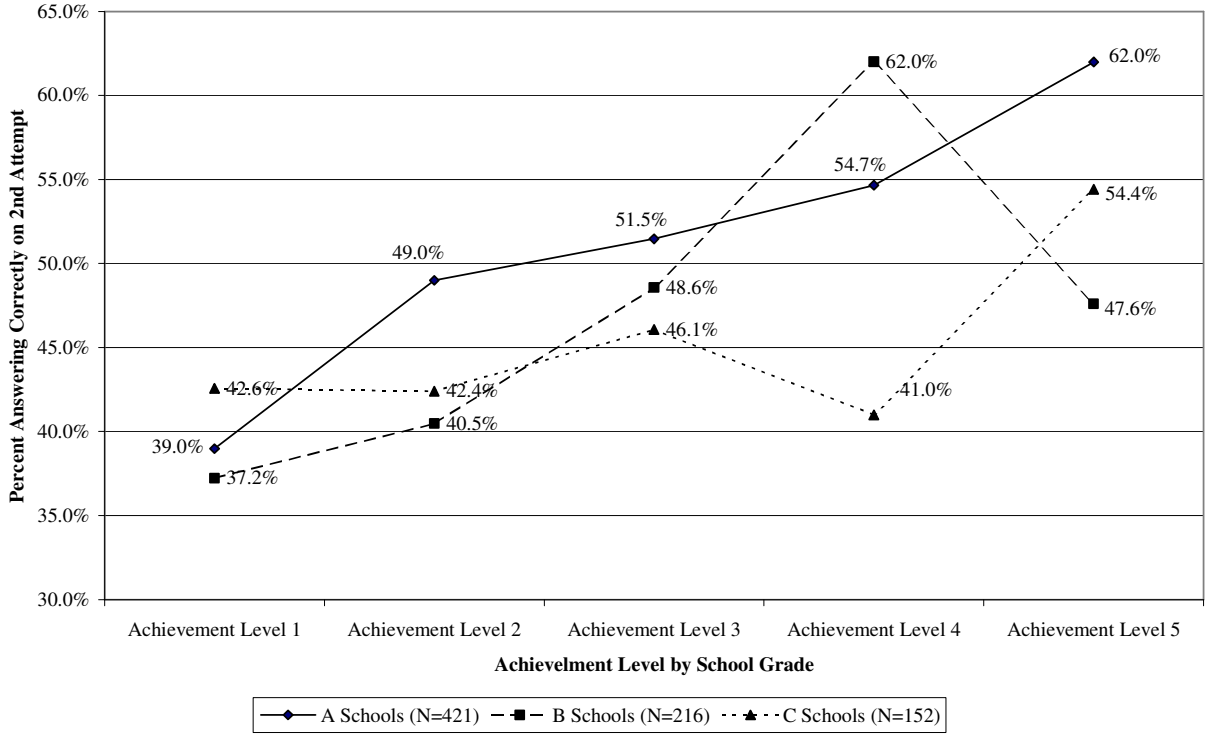


Figure 5. Comparison of Correct Answers on 2nd Attempt Only (P-Value of 2nd Attempt) by Achievement Level and School Grade



Once again, students in the A schools perform better than students in other schools after receiving hint feedback. Surprisingly, though, in the B schools students in Achievement Level 4 outperformed all others on second attempt. (Because there were only 11 students in this group, this might indicate a computer lab setting with a teacher working along with the students.) The results further imply that students at higher levels understand and learn better from the hint feedback. Students in the lowest achievement levels might have difficulties understanding the math problem or comprehending the hint feedback pages due to untapped factors such as reading proficiency.

These two figures demonstrate a relationship between prior performance on the 2001 math FCAT (as indicated by achievement level) and later performance on the FCAT Explorer, in the 2001-2002 school year. Other figures would demonstrate a similar relationship; for example, charting the percentage of incorrect answers would show a downward sloping line with the highest achievement level getting the fewest incorrect answers. Another way to demonstrate these relationships is to map them out in a correlation table. Table 4 shows the set of correlations between the FCAT Explorer practice and feedback indicators and the 2002 math FCAT scores.

Correlation of FCAT Explorer Effects Model with 2002 Math FCAT Scores

The correlation table above indicates that most of the practice and feedback indicators have significant relationships with the 2002 math FCAT scores. The strongest correlation in the table, 0.839, is between the 2001 and 2002 math FCAT scores, which demonstrates the predictive power of past performance. At 0.510, the knowledge indicator, Percent Correct at 1st Attempt, has the strongest correlation with the 2002 math FCAT score among the 5GM indicators. This positive relationship is followed by the learning indicator, Percent Correct on 2nd Attempt Only, which is also significantly correlated at 0.274. The percent of Incorrect and Incomplete answers are negatively correlated, as might be expected, since these are indicators of a lack of learning from the hint feedback pages or giving up on attempting the 2nd answer. The Percent of Correct 2nd Answers Across all Items, and the number of Hint Feedback pages are also negatively correlated with the 2002 math FCAT scores. This makes sense if one considers that the hint feedback pages and second attempt occur only after the student has answered incorrectly the first time, so as a student answers more items correctly, the percentage of 2nd attempts will be reduced; thus, fewer 2nd attempts correlate to better FCAT scores. Two indicators have no

relation to the 2002 math FCAT scores, Total Answers and Number of Correct Answer Pages, suggesting that quantity alone is not significant in this model. These indicators are dropped from the model in the next step of testing.

Table 4. Correlations Between Performance Indicators in the FCAT Explorer Effect Model and 2001 FCAT Score and Change in FCAT Score Between 2001 and 2002

FCAT Explorer Effects Model Indicators	2002 FCAT Scale Score for Math	
	Pearson r	Sig.
Practice Indicators		
Total Items Attempted	0.021	0.521
Percent of Correct 1st Answers (P-Value of 1st Attempt)	0.510	0.000
Percent of Correct Answers on 2nd Attempt of All Items	-0.143	0.000
Percent of Incorrect Answers	-0.353	0.000
Percent of Incomplete Answers	-0.223	0.000
Feedback Indicators		
How Many Times Student Received Hint Feedback	-0.206	0.000
Number of Correct Answer Pages Presented	0.034	0.305
Percent of Correct 2nd Answers on 2nd Attempt Only (P-Value of 2nd Attempt)	0.274	0.000
2001 Math FCAT Score		
2001 FCAT Scale Score for Math	0.839	0.000

Multiple Regression Test of FCAT Explorer Effects Model

The final test of the FCAT Explorer effects model is to use a linear regression statistic to determine how much of the variance in the 2002 math FCAT score can be attributed to the indicators in the model. The output of the multiple regression model is shown in Table 5, which lists the six practice and feedback indicators in the effects model and their explanatory relation to the 2002 math FCAT scores. Of the six, only three show a significant relationship, the percent of correct answers on 1st attempt, the correct answer on 2nd attempt only and the percent of incorrect answers. The other indicators do not appear to affect math FCAT performance. The effects model shown in Table 5 is significant at $p < 0.000$, based on an ANOVA calculation, and has an

adjusted R^2 of 0.311, which indicates that about 31% of the variance in the 2002 math FCAT scores is explained by the indicators in the effect model.

Table 5. Multiple Regression of Full FCAT Explorer Effects Model on 2002 FCAT Scores

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.562	.316	.311	40.511

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	688555.103	6	114759.184	69.925	.000
	Residual	1491827.198	909	1641.174		
	Total	2180382.300	915			

b. Dependent Variable: FCATSS02 FCAT Scale Score for Math 2002

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	284.722	16.168		17.610	.000
	CORR1STP Percent of Correct 1st Answers (P-Value of 1st Attempt)	99.392	13.915	.403	7.143	.000
	PCOR2ND Percent of Correct Answers on 2nd Attempt of All Answers Attempted	-43.313	26.106	-.105	-1.659	.097
	INCORRP Percentage of Incorrect Answers	-25.202	10.288	-.123	-2.450	.014
	INCOMPLP Percent Incomplete Answers	22.634	16.831	.053	1.345	.179
	FDBKPGS Total Number of Hint and Correct Answer Feedback Pages Shown	4.708E-03	.019	.007	.248	.804
	CORR2NDP Percent of Correct 2nd Answers on 2nd Attempt (P-Value of 2nd Attempt)	41.001	10.541	.202	3.890	.000

a. Dependent Variable: FCATSS02 FCAT Scale Score for Math 2002

The indicator for knowledge, Percent Correct on 1st Attempt, is the strongest predictor of math FCAT performance, with a t score of 7.143, a beta weight of 0.403 and a significance level of $p < 0.000$. This is verification that answering the item correctly the first time is the best predictor of FCAT success. With such a strong explanatory relationship, it also suggests that the FCAT Explorer 5GM could be used as a diagnostic tool to indicate which students would do well and which would not do well on the math FCAT.

The second strongly significant indicator is the Percent Correct on 2nd Attempt Only, which has a t score of 3.89, a beta weight of 0.202 and is significant at $p < 0.000$. This is an important finding because it points to the value of the hint feedback as a learning opportunity for students. Students who were able to incorporate the hints into their knowledge base and then answer the math problem also do better on the math FCAT than students who did not. This last is shown by the third significant indicator in the model, the Percent of Incorrect Answers. Its relationship with the 2002 math FCAT scores is negative, as one might expect, with t score of –2.45, a beta weight of -0.123 and a significance level of $p < 0.014$. This last, negative, relationship has less explanatory power, and exists, really, as a "shadow" indicator of the 1st and 2nd correct. Nonetheless, this finding points out that the fewer incorrect answers a student has, the better he or she will do on the math FCAT. None of the other indicators in the effects model has a relationship to the 2002 math FCAT, and thus can be ignored.

From the set of figures and tables shown in this section, it is clear that both practice and feedback components of the FCAT Explorer 5GM program have an effect on math FCAT scores. While there are clear differences in the performance of students on the FCAT Explorer, by achievement level and school grade, the similarities are also striking. The ability to answer correctly is clearly related to the achievement level of the student, based on how well the student did on the 2001 math FCAT. Because past performance should predict future performance, it comes as no surprise that students who did well on the 2001 FCAT in 2001 should do well on the FCAT Explorer and on the 2002 math FCATs, as indicated by their placement in achievement levels. The question could then be raised that the relation between using the FCAT Explorer and the 2002 FCAT scores is an artifact of the relationship between performance on the 2001 math FCAT and the 2002 math FCAT. To test this possibility, a final multiple regression was run comparing the effects of the 2001 math FCAT Score, Percent Correct on 1st Attempt and Percent Correct on 2nd Attempt Only on the 2002 math FCAT Scores. Table 6 shows the outcome of

statistical test. By including the 2001 math FCAT Scores the effects model increases in explanatory strength with an adjusted R^2 of 0.689. Each of the variables in the model are significantly related to the 2002 math FCAT Scores, confirming that use of the FCAT Explorer 5GM program does have an effect, over and above the students' previous FCAT performance.

Table 6. Comparison of the Effects of 2001 Math FCAT Scores and FCAT Explorer 5GM Activity on the 2002 Math FCAT Scores

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.831	.690	.689	27.428

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1414217.075	3	471405.692	626.613	.000
	Residual	634947.905	844	752.308		
	Total	2049164.980	847			

b. Dependent Variable: FCATSS02 FCAT Scale Score for Math 2002

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	123.783	5.195		23.829	.000
	CORR1STP Percent of Correct 1st Answers (P-Value of 1st Attempt)	33.695	5.424	.137	6.212	.000
	CORR2NDP Percent of Correct 2nd Answers on 2nd Attempt Only (P-Value of 2nd Attempt)	15.077	4.018	.074	3.753	.000
	FCATSS01 FCAT Scale Score for Math 2001	.598	.018	.734	32.814	.000

a. Dependent Variable: FCATSS02 FCAT Scale Score for Math 2002

Students' 2001 math FCAT scores are clearly the best predictor of 2002 math FCAT performance, with a t score of 32.814, a beta weight of 0.734, and is significant at $p < 0.000$. The Percent of Correct 1st Answers is also significantly related to the 2002 math FCAT score, though it has less weight with the external factor of the 2001 math FCAT score entered into the model. It is a predictor with a t score of 6.212, a beta weight of 0.137 and is significant at $p < 0.000$. This

finding means that whether a student did well or poorly on the 2001 math FCAT, practicing on the FCAT Explorer 5GM helped his or her performance on the 2002 math FCAT. The Percent Correct on 2nd Attempt Only variable, which indicates student learning, is also significantly related to the 2002 math FCAT score, again to a lesser degree. Its t score is 3.753, with a beta weight of 0.074; but it is still significant at $p < 0.000$. This final test of the model only confirms the positive effect of using the FCAT Explorer 5GM on students' math FCAT performance.

Conclusions

This study has tried to answer a practical question: Does using the FCAT Explorer 5GM during the 2001-2002 school year have a positive effect on student performance on the 2002 math FCAT? From the data analysis it is clear that both the practice and feedback components in the FCAT Explorer have an impact on students' performance on the math FCAT.

The analysis demonstrates that student performance on the FCAT Explorer varied with the achievement level of the student, regardless of school grade. The lower the achievement levels did not perform as well as the higher achievement levels in either answering items correctly at 1st attempt on the FCAT or on 2nd attempt. The first attempt at answering a math item is indicative of the students' knowledge of the subject. Even in the lowest achievement levels students could answer correctly on first attempt better than 40% of the time; the highest achievement level students answered correctly about 70% of the time. The students who did not answer the item correctly then receive a set of feedback pages designed to explain their mistake and offer some hints for them to answer correctly. Again, students in the lowest achievement level did not perform as well as the higher achievement levels, but another 40% of them answered correctly the 2nd time around. This finding indicates that students who are struggling with their math studies can benefit from a program like the FCAT Explorer.

The test of the FCAT Explorer effects model also indicates the primacy of answering correctly both on 1st attempt and on 2nd attempt, for all students in the sample. These two indicators are significant factors that explain about 20% of the variance in the 2002 math FCAT scores when compared to the students' performance on the 2001 math FCAT. This suggests that practice followed by learning feedback and more practice enhances learning and thus has a positive effect on the students' performance on the FCAT. None of the other factors in the FCAT

Explorer 5GM effects model has the same impact, neither the item counts nor the other percentage indicators.

In summary, we have found that the first research hypothesis proposing that the number of practice items will have an influence on a student's FCAT performance is not supported. This suggests that the quantity of items practiced is not as important. The second research hypothesis, that correct answers will affect FCAT performance is supported. Regardless of the number of items or math problems, how well a student answers those items is the important factor here. It is not the quantity of items but the quality of the responses that affects students learning. The third research hypothesis, that students who do not do well on the FCAT Explorer will likewise not do well on the FCAT is supported with a relatively weak, but statistically significant finding that a student's percentage of incorrect answers is negatively related to the math FCAT scores. Finally, the last research hypotheses that reading more feedback pages will help a student perform better on the FCAT is not supported in terms of the number of pages. The outcome of reading the feedback pages and answering correctly on 2nd attempt is discussed in the second research hypothesis.

This study points out an important relationship between the use of an educational software program and performance on the math FCAT. From the analysis is appears clear that using the FCAT Explorer 5GM does contribute to students' performing well on the FCAT. A final summary of the data shows that 32% of all students increased their achievement level between 2001 and 2002, across all school grades. From this evidence one can safely say that the FCAT Explorer 5GM plays a significant role in helping 5th graders prepare for the math FCAT.

References

Babbitt, B. C. and Miller, S. P., (1996). Using hypermedia to improve the mathematics problem-solving skills of students with learning disabilities. *Journal of Learning Disabilities,* v 29, p. 391-401.

Ball, D. L., & Wilson, S.M. (1990). Knowing the subject and learning to teach it: Examining assumptions about becoming a mathematics teacher (Research report 90-7). East Lansing, MI: National Center for Research on Teacher Learning, Michigan State University.

Bloom, B. S. (1956). Taxonomy of educational objectives: The classification of educational goals, Handbook I: Cognitive domain. New York: Longman & Green.

Cooper, S. B. (1998). Instructor-created computer tutorials for students in an elementary mathematics education course. *Journal of Computing in Childhood Education,* v 9 no1, p. 93-101.

Dempsey, J. and Sales, G. (1993) (Edited). Interactive Instruction and Feedback. Englewood Cliffs: Educational Technology Publications.

Dick, W., & Carey, L. (1990). The systematic design of instruction. Glenview, IL: Scott, Foresman.

Fahy, P. J. (2000). Achieving quality with online teaching technologies. Paper presented at Quality learning 2000 Inaugural International Symposium, Calgary, Alberta, March 2000. (ERIC, ED 445 197.

Freitag, E. T., & Sullivan, H. J. (1995). Matching learner preference to amount of instruction: An alternative form of learner control. *Educational Technology Research and Development*, 43(2), 5-14.

Gagne, R. M. (1987). The conditions of learning (4th ed.). New York: Holt, Rinehart, and Winston.

Goldengerg, E. P., & Cuoco, A. (1996). What is dynamic geometry? In R. Lehler & D. Chazan (Eds.), Designing learning environments for developing understanding of geometry and space. Hillsdale, NJ: Erlbaum.

Golding, J, 2003. Testing, testing: The place of assessment in the mathematics classroom. Mathematics in School v. 32 no3, p. 12-13

Goyne, June S., McDonough, Sharon K., Padgett, Dara D., 2000. Practical guidelines for evaluating educational software. The Clearing House, v. 73 no6, p. 345-8

Hannafin, R. D., & Sullivan, H. J. (1996). Learner preferences and learner control over amount of instruction. *Journal of Educational Psychology*, 88, 162- 173.

Hannafin, R. D.; Scott, B. N. (1998). Identifying critical learner traits in a dynamic computer-based geometry program. *The Journal of Educational Research* (Washington, D.C.), v. 92 no1, p. 3-12.

Hawkins, J. (1993). Technology and the organization of schooling. Communications of the ACM, 36 (5), 30 – 34.

Herrington, J. & Oliver, R. (1999). Using situated learning and multimedia to investigate higher-order thinking. *Journal of Interactive Learning Research,* v 10 no1, p. 3-24.

Hopson, M. H., Simms, R. L., & Knezek, G. A. (2001-2002). Using a technology-enriched environment to improve higher-order thinking skills. *Journal of Research on Technology in Education, 34*(2), p. 109-119.

Keller, J. M. (1987). Development and use of the ARCS model of instructional Design. *Journal of Instructional Development, 10*(3), 2-10.

Khine, M. S. (1996). The interaction of cognitive styles with varying levels of feedback in multimedia presentation. *International Journal of Instructional Media,* v. 23 no3, p. 229-37.

Kim M. K.; Sharp, J. M. (2000). Investigating and measuring preservice elementary mathematics teachers' decision about lesson planning after experiencing technologically enhanced methods instruction. *The Journal of Computers in Mathematics and Science Teaching,* v.19 no4, p. 317-38.

Kulhavy, R.W. and Wager, W. (1993). Feedback in programmed instruction: Historical context and implications for practice. In Dempsey, J. and Sales, G. (Edited). Interactive Instruction and Feedback. Englewood Cliffs: Educational Technology Publications.

Naime-Diefenbach, B. (1991). Validation of Attention and Confidence as independent components of the ARCS motivational model. Unpublished doctoral dissertation. Florida State University, Tallahassee, Florida.

Naime-Diefenbach, B. and Sullivan, C.B. (2001). Measuring the effects of a web-based practice program on the FCAT math results of 5th graders. Paper presented at the Florida Educational Research Association Annual Conference, November 7 - 9, 2001, Marco Island, Florida.

Nicaise, M., (1997). Computer-supported apprenticeships in math and science. *The Journal of Computers in Mathematics and Science teaching*, v.16 no4, p.443-65.

Paolucci, R. (1998). The effects of cognitive style and knowledge structure on performance using hypermedia-learning system. *Journal of Educational Multimedia and Hypermedia*, v. 7 no2-3, p. 123-50.

Russell, T. (1997). Teaching teachers: How I teach IS the message. In J. Loughran & T. Russell (Eds.). Teaching about teaching: Purpose, passion and pedagogy in teacher education. London: Falmer.

Sasser, J. E. (1990-1991). The effect of using computer tutorials as homework assignments on the mathematics achievement of elementary education majors. *Journal of Computers in Mathematics and Science Teaching*, 10(2), 95-102.

Schank, R.C. (1993). Learning via multimedia computers. Communications of the ACM, 36(1), 54-56.

Schifter, D. (1997). Learning mathematics for teaching: Lessons in/from the domain of fractions. Newton, MA: Education Development Center, Inc. (ERIC Document Reproduction Service No. ED 412 122).

Scholz, J. M. (1995). Professional development for mid-level mathematics. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA. (ERIC Document Reproduction Service No. ED 395 820).

Schoenfeld, A. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D.A. Growa (Ed.), Handbook of Research on mathematics teaching and learning (pp. 334-370). New York: Macmillan.

Shute, V. J. (1993). A comparison of learning environments. In S. P. Lajoie & J. Derry (Eds.), Computers as Cognitive tools (pp. 46-64). Hillsdale, NJ: Erlbaum.

Snider, R. C. (1992). The machine in the classroom. Phi Delta Kappan, 74 (4), 316 – 323.

Torrez, N. (2000). Developing Culturally Consonant Curriculum Using the Technology of the New Millennium. Paper presented at the Annual Meeting of the American Association of Colleges for Teacher Education (52nd, Chicago, IL, February 26 – 29, 2000).

Verzoni, K. A. (Oct. 1997). Turning students into problem solvers. *Mathematics Teaching in the Middle School*, v.3, p.102-7.

Appendix A. Distribution of Students in Achievement Levels by School Grade

Achievement Level of Student	2001 School Grade A (n=458)		2001 School Grade B (n=227)		2001 School Grade C (n=168)		2001 School Grade N (n=63)		Total	Average % Ach Levels
	n	%	n	%	n	%	n	%	n	%
2001 Achievement Level 1	27	5.9%	20	8.8%	57	33.9%	13	20.6%	117	12.8%
2001 Achievement Level 2	69	15.1%	52	22.9%	43	25.6%	16	25.4%	180	19.7%
2001 Achievement Level 3	127	27.7%	73	32.2%	39	23.2%	17	27.0%	256	27.9%
2001 Achievement Level 4	132	28.8%	48	21.1%	11	6.5%	11	17.5%	202	22.1%
2001 Achievement Level 5	66	14.4%	23	10.1%	2	1.2%	2	3.2%	93	10.2%
Missing	37	8.1%	11	4.8%	16	9.5%	4	6.3%	68	7.4%
Total	458	100%	227	100%	168	100%	63	100%	916	100.0%

Appendix B. Comparison of Student Averages on FCAT Explorer Effects Model Indicators, by 2001 School Grade and Achievement Level

Ach.Level 2001 (1-5)	Average Items Attempted	Average Percent of Correct 1st Answers	Average Percent of Correct Answers on 2nd Attempt, All Answers	Average Percent of Total Correct Answers	Average Percentage of Incorrect Answers	Average Percent of Incomplete Answers	Average Times Student Received Hint Feedback	Average Percent of Correct 2nd Answers on 2nd Attempt Only
2001 A Schools								
Ach. Level 1	72.5	40.8%	21.7%	63.6%	61.0%	4.3%	43.3	39.0%
Ach. Level 2	48.9	50.7%	22.4%	74.2%	51.0%	4.4%	24.6	49.0%
Ach. Level 3	47.7	59.9%	19.8%	80.9%	45.4%	4.9%	19.6	51.5%
Ach. Level 4	49.4	67.4%	17.3%	85.1%	41.6%	4.3%	16.6	54.7%
Ach. Level 5	53.7	75.0%	16.4%	91.7%	30.4%	1.7%	13.9	62.0%
2001 B Schools								
Ach. Level 1	38.5	46.3%	19.4%	67.0%	57.8%	5.3%	22.4	37.2%
Ach. Level 2	29.2	50.5%	19.8%	71.4%	51.8%	7.6%	14.1	40.5%
Ach. Level 3	23.9	57.1%	19.0%	77.0%	44.6%	7.7%	9.5	48.6%
Ach. Level 4	25.2	61.5%	23.8%	86.7%	35.9%	5.1%	8.8	62.0%
Ach. Level 5	35.9	66.1%	12.4%	78.7%	35.0%	15.1%	8.2	47.6%
2001 C Schools								
Ach. Level 1	35.3	40.8%	22.6%	65.0%	55.7%	9.2%	19.3	42.6%
Ach. Level 2	34.3	46.6%	20.6%	68.7%	57.6%	6.9%	17.6	42.4%
Ach. Level 3	42.6	51.8%	20.5%	73.1%	53.9%	4.2%	21.0	46.1%
Ach. Level 4	58.7	67.6%	14.0%	81.9%	49.9%	3.2%	21.5	41.0%
Ach. Level 5	32.5	70.6%	16.1%	87.0%	45.6%	2.5%	10.5	54.4%