5-2003

An Investigation of the Impact of Integrated Learning System Use on Mathematics Achievement of Elementary Students.

Vicki Carpenter Kirk
East Tennessee State University

Investigation of the Impact of Integrated Learning System Use on Mathematics Achievement of Elementary Students

---

A dissertation
presented to
the faculty of the Department of Educational Administration and Policy Analysis
East Tennessee State University

In partial fulfillment
of the requirements for the degree
Doctor in Education

---

by
Vicki Carpenter Kirk
May 2003

Dr. Louise MacKay, Chair
Dr. Russell Mays
Dr. Russell West
Dr. Cecil Blankenship

Keywords: educational technology, mathematics achievement, integrated learning system, computer-assisted instruction

ABSTRACT

An investigation of the Impact of Integrated Learning System Use

on Mathematics Achievement of Elementary Students

By

Vicki Carpenter Kirk

The purpose of this study was to explore the relationship between time spent on an integrated learning system (ILS) entitled SuccessMaker® as a supplement to traditional mathematics instruction on achievement as measured by standardized achievement tests of elementary students. The variables of grade-level, ability level, and gender were also considered. The population consisted of students who were second-, third-, and fourth-graders during the 1997-98 school year. Data were gathered that covered the three-year period beginning in 1997 and ending in 2000. The population consisted of 348 students who participated in Computer Curriculum Corporation© mathematics instruction and who completed the Terra Nova in 1997-98, 1998-99, and 1999-2000. Analysis of Variance was used to identify any relationship between variables.

The study's investigation of the relationship between ILS use and mathematics achievement could assist educators in planning for use of technology as a supplement to traditional instruction. While the information gleaned is specifically beneficial to Greeneville City Schools, other school systems seeking information on the relationship between ILS use and achievement will find this study constructive, especially when viewed in conjunction with the existing body of literature.

Findings in this study were mixed. ILS use was associated with positive effects, negative effects, and no effects. It was noted that negative effects occurred during the year with the lowest overall usage. No interaction effects were found in any of the models, indicating that the ILS did not have differing effects for boys or girls or for students of varying ability levels. Positive effects of the ILS, Math Concepts and Skills (MCS), on math composite scale scores were noted at grades two and three, while students at grades four, five, and six were either unaffected or negatively affected by the use of MCS. Math Investigations (MI), although used on a very limited basis during the course of this study, had a positive effect overall on math composite scale scores. Clearly, when math gain was the dependent variable, there were no effects demonstrated by use of MCS or MI.

DEDICATION

This work would not have been possible without the loving support of my family. This book is dedicated to them. My husband, Michael, and my sons, Ben and Robert, have been loving, supportive, and patient. Thank you for putting up with the complaints, the time requirements, and the mess. To my parents, my first teachers and biggest fans, I am eternally grateful. I love you all dearly;
thank you for your confidence in me.

ACKNOWLEDGEMENTS

I would like to acknowledge the tireless assistance of my graduate committee members,

Dr. Louise MacKay, Chair

Dr. Russell Mays

Dr. Russell West

Dr. Cecil Blankenship

I would also like to acknowledge the encouragement and guidance of my dear friend, Dr. Linda Stroud. She provided the impetus for embarking on this journey and was the "keeper of the timeline" as well as "provider of practical advice".

Finally, I would like to acknowledge the assistance of Dr. Susan Twaddle, whose expertise and assistance were invaluable to the completion of this dissertation.

CONTENTS

Chapter Page

Chapter Page

Chapter Page

LIST OF TABLES

Table Page

CHAPTER 1
INTRODUCTION

As we embrace the new millennium, U.S. schools are arming themselves with technology to prepare students for a rapidly changing future. Indeed, this effort is a national initiative. In his State of the Union Address on January 23, 1996, President Clinton called for the commitment of resources necessary to provide all students with access to high quality technology and the information superhighway (US Department of Education, 1997). At the same time, the White House announced the following technology goals:

1. All teachers in the nation will have the training and support they need to help students learn to use the information superhighway.
2. All teachers and students will have modern multimedia computers in their classrooms.
3. Every classroom will be connected to the information superhighway.
4. Effective software and on-line learning resources will be an integral part of every school's curriculum. (Coley, Cradler, & Engel, 1997, p. 8)

It seems that significant progress is being made toward these goals. In 1997 there were 4.4 million computers in America's classrooms with a ratio of 10 students per computer (Coley et al.). The computer usage rate of students in grade 4 increased from 61% in 1996 to 83% in 2000 (Braswell, Lutkus, Grigg, Santapau, Tay-Lim, & Johnson, 2001).

The president budgeted $2 billion from 1992 – 1997 for the Technology Literacy Fund to help states implement their educational technology plans (US Department of Education). Archer (1998) estimated the nation's annual investment in educational technology at more than $5 billion. Yet another estimate, proposed by McKinsey & Company's Classroom Model (US Department of Education, 1997) sets the investment at approximately $11 billion per year over 10 years. This would result in the realization of President Clinton's vision of one computer for every five students, access to the information superhighway, and appropriate software in every classroom (US Department of Education). Greeneville City Schools has made such an investment. Over a period of three years $2.2 million has been spent to provide each classroom with five computers; each is connected to the worldwide web and is equipped with a number of grade-level appropriate courseware packages. The primary educational courseware used by Greeneville City Schools is Computer Curriculum Corporations SuccessMaker®, in integrated learning system (*Courseware Description*, 1996).

> Financial investment of this magnitude is usually accompanied by significant accountability. Parents and teachers, school boards and administrators, governors and state legislatures, and Congress all want to know if the nation's investment in technology is providing a return in student achievement. Indeed, if resources are to be expended on technology, it is becoming a political, economic, and public policy necessity to demonstrate its vital effectiveness. (McNabb, Hawkes, & Rouk, 1999, p. 1)

Honey, Culp, and Carrigg (1999) state,

> A key recommendation growing out of the President's Committee of advisors on Science and Technology is the need for large-scale, longitudinal studies that examine the consequences of technology use in school settings in concert with a broad range of factors. (p. 3).

This study focused on Greeneville City Schools' use of an integrated learning system to supplement mathematics instruction. Data gathered over a three-year period were examined to determine if the infusion of dollars in the form of computer hardware, software, and training has made a measurable impact on student learning. It is hoped that this study will provide useful information in identifying patterns and trends in the effective use of technology that will result in optimum benefits to students.

Statement of the Problem

The purpose of this study was to determine if a relationship existed between time spent on an integrated learning system entitled SuccessMaker® as a supplement to traditional mathematics instruction and mathematics achievement as measured by standardized achievement tests of elementary students at various levels of achievement. The variables of grade-level and gender were also considered.

Research Questions

The following questions related to the use of an integrated learning system (ILS) as a supplement to traditional mathematics instruction were addressed:

1. How did the population's achievement scores and gain scores compare with those of the nation?
2. How much time did students spend using Math Concepts and Skills (MCS)?
3. Has change occurred in cohorts' use of the ILS over the three-year period?
4. To what extent do boys and girls use MCS?
5. To what extent do students in different ability groups use MCS?
6. Did cohort members exhibit a difference in cumulative use of MCS?
7. How much time did students spend using Math Investigations (MI)?
8. To what extent do boys and girls use MI?
9. To what extent do students in different ability groups use MI?
10. Does MCS use have an effect on mathematics achievement as measured by math composite scale scores?
11. Does two years of cumulative use of MCS have an effect on mathematics achievement as measured by math composite scale scores?
12. Does three years of cumulative use of MCS have an effect on mathematics achievement as measured by math composite scale scores?
13. Does MI use have an effect on mathematics achievement as measured by math composite scale scores?
14. Does MCS or MI use affect math achievement as measured by math composite scale scores based upon gender?
15. Does MCS use have an effect on mathematics achievement as measured by math gain?
16. Does MI use have an effect on mathematics achievement as measured by math gain?
17. Does MCS use have varying affects on mathematics gain with regard to gender or ability level?

18. Does MI use have varying affects on mathematics gain with regard to gender or ability level?

Significance of the Study

In this age of accountability, educators must ensure that the strategies and interventions they employ are effective. Weaver (2000) stated,

> In light of the literature that indicates both the prevalence of instructional technology and the scarcity of detailed assessment of its effect on learning, it is evident that the use of computers in education is at a critical juncture in its history. (p. 131).

Student achievement is being scrutinized. McNabb, Hawkes, and Rouk (1999) conceded that, “Standardized test scores have become the accepted measure with which policy makers and the public gauge the benefits of educational investments” (p. 5). Under this kind of pressure, teachers and administrators are searching for better, more efficient methods of delivering instruction. Use of technology for this purpose appears promising. Niemiec, Blackwell, & Walberg (1986), in a quantitative synthesis of 48 studies, estimated an average computer-assisted instruction (CAI) effect of 0.45, a significant positive effect. More recent research (Mann & Schafer, 1997; Mann, Shakeshaft, Becker, & Kottkamp, 1999; Schlago-Schirm, 1995; Weaver, 2000; Wenglinsky, 1998) also demonstrated a positive relationship between CAI and student achievement.

Technology should not, however, be viewed as a panacea. Indiscriminate, inappropriate use of technology could be harmful. In fact, research demonstrating both positive and negative effects can be cited. While Van Dusen and Worthen (1994) have found that computer-based integrated learning systems can be very effective, DeVaney (1996) reported that student achievement in geometry was negatively impacted by frequent use of computers and the use of computers for drill and practice.

Starr (1996) counseled that efforts to improve education with better technology “reflect a persistent infatuation with technological fixes for deeply rooted social problems” (p. 50). He expressed the hope that more didactic forms of computer-assisted instruction could be used to prepare students for standardized tests, freeing up remaining instructional time for more progressive teaching strategies. The debate regarding effective use of instructional learning systems (ILSs) is legitimate and necessary. Teachers have limited time to teach increasingly complex curricula. They want assurance that the investment of precious instruction time in ILS use is a wise one.

Greeneville City Schools, in addition to providing state-of-the-art technology for every classroom, has recently placed emphasis on improving math achievement test scores. It is hoped that the individualized plan enabled by the use of the integrated learning system will achieve the ambitious goal of meeting each student at his or her present ability level and helping him or her to progress efficiently and effectively.

Mann (1999) stated that technology is no different from any other pedagogy in that it only works for some students, in some topics, and under some conditions; it must be used judiciously. Without adequate research, judicious use of technology becomes a trial and error process. Mann decried the fact that a generation of implementation has failed to provide good measures of children's exposure or even of teachers' use of programs. He emphasized a need for much more attention to elapsed time, exposure effects, and dosage effects. These issues were addressed in this study.

This study provided useful information about the use of an integrated learning system. Information gathered over a period of three years offered insight into the effect that time spent using an ILS had for children of various achievement levels for children at different grade levels and for boys as well as girls. Data collected over a three-year period make the results of this study particularly useful. While the information gleaned is specifically beneficial to Greeneville City Schools, other schools and school systems seeking information on the relationship between CAI use and achievement will find this study constructive, especially when viewed in conjunction with the existing body of literature.

Definitions

Achievement Test: "An assessment that measures a student's currently acquired knowledge and skills in one or more of the content areas common to most school curricula (for example, reading, language arts, mathematics, science, and social studies)" (*Test Coordinator's Handbook*, 1997, p. 42).

Average-Achieving Students: Achievement groups were determined using sample statistics from the initial test administration for each cohort group. Students performing within one standard deviation of the mean were considered to be average. For the second grade cohort the range was 568 – 603; for the third grade cohort it was 604 – 635, and for the fourth grade cohort the range was 632 – 665.

Computer-Assisted Instruction (CAI): "An instructional technique based on the two-way interaction of a learner and a computer with the objective of human learning and retention" (Division of Educational Sciences, 1987, p. 30).

Courseware: "The actual instructional material, including both content and technique, installed on a CAI system" (Division of Educational Sciences, p. 34). This differs from software, which is the actual program.

Gain Score: A gain score is the difference between scale scores from one year to the next (*Interpreting Tests*, www.quia.com).

High-Achieving Students: Achievement groups were determined using sample statistics from the initial test administration for each cohort group. Those students who scored more than one half standard deviation above the mean were considered to be high-achieving. For grade the second, third, and fourth grade cohorts those levels were greater than 603, 635, and 665, respectively.

Integrated Learning System (ILS): "Networked comprehensive basic skills software from a single vendor" (Becker, 1992a, p. 1).

Low-Achieving Students: Achievement groups were determined using sample statistics from the initial test administration for each cohort group. Those students who scored more than one half standard deviation below the man were considered to be low-achieving. For the second, third, and fourth grade cohorts those levels were less than 568, 604, and 632, respectively.

Scale Score: "The scale score describes the achievement on a continuum that in most cases spans the complete range of Kindergarten through Grade 12. These scores can range in value form approximately 100 to 900" (*Beyond the Numbers,* 1997, p. 3). "They are units of a single, equal-interval scale that is applied across all levels…regardless of grade or time of year of testing" (p. 48). Scale scores may be mathematically manipulated for statistical purposes.

Delimitations

This study was delimited to the following characteristics, The population consisted of students who were second-, third, or fourth-graders during the 1997-98 school year; these students attended the four elementary schools in the Greeneville City School System. Data were gathered that covered the three-year period beginning in 1997 and ending in 2000. The oldest group entered Greeneville Middle School as sixth-graders during the 1999-2000 school year.

During the three-year period for which data were collected students received both traditional math instruction and computer-assisted instruction. The integrated learning system used the Computer Curriculum Corporation© (CCC) courseware SuccessMaker® and was delivered in each classroom on five personal computers, an instructional model commonly referred to as a distributed setting. Curriculum material was limited to mathematics. Student achievement was measured using the mathematics subtests on the *Terra Nova Achievement Test*, CTBS-5 Edition, Forms I, J, and K (R. Wankel, personal communication, February 20, 2002).

Overview of the Study

This study is organized into five chapters. Chapter 1 provided an introduction, statement of the problem, a listing of research questions and explanation of the significance of the study, pertinent definitions, and delimitations. Chapter 2 presents a review of the literature and is organized into the following sections: historical perspectives; integrated learning systems description, potential, and implementation; research on student achievement and the use of integrated learning systems; methodological problems and suggestions for future research; and summary. Research methodology is detailed in chapter 3. Information is provided on research design, population, student achievement, description of Greeneville City implementation, description of courseware, data collection, and data analysis. Chapter 4 provides information detailing results of the study and is organized into five sections, each of which is associated with one or more of the research questions. A summary of the study is provided in chapter 5, followed by conclusions of the study and recommendations for both practice and further research.

CHAPTER 2
REVIEW OF RELATED LITERATURE

This review of related literature has as its focus the use of computer-assisted instruction (CAI) in elementary schools. The first section relates a historical perspective of the development of CAI. Section two describes integrated learning systems (ILSs), detailing their potential for improving student achievement and providing suggestions for proper implementation.

A review of recent research studies comprises the third section. Descriptions of the research designs, populations, samples, and results of the studies are reported. Individual evaluations and meta-analyses, which are representative of the increasing number of studies being conducted by school districts, individuals, and integrated learning system publishers, are reviewed. The focus is on studies that address the relationship between integrated learning system use and math achievement.

The fourth section highlights methodological problems with many existing studies. Several studies demonstrating dramatically positive results are critiqued showing significant flaws and corrected results. Also included in section four are suggestions for research designs that will improve the quality of ILS research.

Historical Perspective

The quest to use technology to facilitate teaching and learning has been underway since the 19^{th} century. At that time electro-mechanical “teaching machines” were in use. These ancient ancestors of the present-day integrated learning systems were based upon operant-conditioning theory (Becker & Hativa, 1994). “Most of the original ILSs weren’t really much more than systematically organized collections of drill-and-practice software” (Sherry, 1990, p. 118).

Early uses of technology reflect a theory of learning that is individualistic and solitary; a Skinnerian model based on teaching machines and programmed learning. This philosophy assumed that human learning is more psychological than social, that the primary relationship exists between the learner and the content to be mastered (Becker, 1992b).

The first users of computer-assisted instruction (CAI) technology were members of the computer industry who used computer-based instruction as early as the late 1950s. “Educational CAI was an almost natural combination of emerging computer technology and the programmed instruction movement” (Burns & Bozeman, 1981, p. 33).

IBM led the way for other computer companies into the field of instructional computing throughout the 1960s (Burns & Bozeman, 1981). At that time two individualized instruction-oriented, computer-based learning systems were introduced; Programmed Logic for Automatic Teaching Operations (PLATO) and Time-shared Interactive Computer Controlled Information Television (TICCIT). Both were geared for high school, college, and adult students (Becker & Hativa, 1994).One of the most significant CAI models to emerge from this movement was the Stanford Project. This endeavor, led by Patrick Suppes and Richard Atkinson of Stanford University, was among the earliest attempts of CAI use in public education. This small tutorial system in elementary mathematics and language arts was released in 1963 (Becker & Hativa; Burns & Bozeman). The problems these children received were determined by their previous performance, so the lessons were individualized based on student performance. Over time this system demonstrated effectiveness in raising students’ standardized test scores. This system became the licensing agent for other organizations and sparked development of competing systems with the same basic premise (Becker & Hativa).

This Individual Communication (INDICOM) system was a collection of the teacher-authored CAI packages in 11 content areas within the disciplines of mathematics and language arts. This system, launched in 1967, was the first public school CAI project in the Midwest (Burns & Bozeman, 1981).

PLATO was developed in 1960 at the University of Illinois. This program, with the goal of automating individual instruction, developed into a simultaneous system for time-sharing by students. PLATO IV supported several hundred terminals at dispersed locations. Each terminal had access to a central lesson library (Burns & Bozeman, 1981).

Major improvements in CAI were facilitated by innovations in technology during the1980s. This time period saw the development of stand-alone microcomputers replacing terminals that had served only as keyboard input and video screen output. During the mid to late 1980s sophisticated color graphics were developed that enhanced existing and developing programs. The development of new technologies such as CD-ROM for storing and distributing software made the use of bigger, better, faster, and more sophisticated software commonplace. There has been recent expansion of ILS software beyond simple, skill-oriented presentation to include "tool" software such as word processors, spreadsheets, and information databases (Becker & Hativa, 1994).

"Corporate involvement in CAI development has traditionally taken the form of hardware/software contributions" (Burns & Bozeman, 1981). One example is Computer Curriculum Corporation© (CCC), a direct outgrowth of the Stanford endeavor. CCC offers a variety of sequentially developed CAI drill-and-practice courseware in the areas of mathematics, reading, and language arts (Burns & Bozeman, 1981).

Integrated Learning System Description, Potential, and Implementation

Packages such as those offered by CCC are termed integrated learning systems. Becker (1992a) described integrated learning systems (ILSs) as software provided by a single vendor and housed on a central server. This software contains instruction and practice problems covering a multiple-year curriculum sequence. Lessons are provided for each student based on continuous assessment of that student's previous performance. Sherry (1990) stated that ILSs use networked systems of multiple microcomputers and link lessons to an accepted standard curriculum. Robinson (1991) offered another explanation including the fact that the personal computers (PCs) are linked together in a local area network (LAN) with management systems and instructional software. There is a powerful main computer, a fileserver, where the instructional software, management system, networking software, and student files reside.

Educational software can be classified into four categories. Three of these teach content: 1. drill-and-practice, 2. simulations, and 3. tutorials; the fourth is tools for writing, designing, or creating (Vargas, 1986). Drill-and-practice exercises were designed to increase the speed and accuracy of a skill. Research has demonstrated that time limits on drill-and-practice exercises improved progress over time. Speed is a necessary component of competence in any field, and the use of technology can help students rise above the level of mechanics. Computers provide an ideal medium for providing drill and practice (Vargas, 1986).

As another type of educational software, simulations are useful because they require students to make decisions similar to those required in a real situation. Students may become frustrated initially because simulations do not provide small, sequenced learning steps. Some students may play around without analyzing what they are doing; therefore, teacher planning and supervision are critical when using simulations (Vargas, 1986).

The third category of educational software noted by Vargas (1986) is tutorials, an older, less-used methodology. Tutorials are usually comprised of several screens of text followed by a quiz. They are traditional workbooks on the screen.

Ornstein (1992) stated that computers could be helpful through three learning phases. Through the stage of acquisition, computers can be used to generate and retrieve information. During the transformation phase, self-contained units or modules can be used as instructional supplements to test knowledge, understanding, and application of skills and concepts. Lastly, during evaluation, the computer can provide rapid and accurate feedback. By employing all three phases, ILSs can individualize learning, enabling each student to progress at his or her own pace (Mageau, 1991). Due to this capability, many educators are optimistic about the use of ILSs to improve student achievement.

Several reasons why ILSs should be successful were stated in *Electronic Learning* (Bracey, 1991): (a) Computer activities are motivational for students, (b) the networked system is convenient for teachers, (c) ILSs are adaptable to students' needs, (d) ILSs can be used diagnostically, and (e) using software from a single vendor is less confusing than using software from several vendors. Becker (1992a) provided a similar list of reasons ILSs should be effective in helping to improve student academic performance, particularly in basic mathematics and language skills: (a) Students generally enjoy working at the computer; therefore, computer-based instruction is motivating, (b) because ILSs are networked, many logistical problems are solved, (c) ILSs provide centralized management (allowing teachers to target specific problems for specific students) and individualized skill practice (providing a faster pace for able students and more review for slower students), (d) tutorial capabilities can provide direct instruction where students need it, and (e) because one vendor supplies all the software, there is a consistent user interface which allows students to concentrate on the task at hand rather than figuring out how each program works.

Sherry (1990) stated that the major strength of the ILS is its management system and that “many of the ILSs have evolved into powerful educational tools whose full potential has yet to be tapped” (p. 118). Many school systems are literally buying into this belief. It was reported in 2000 that computers were widely available in 83% of 4th grade classrooms, in 52% of 8th grade classrooms, and in 43% of 12th grade classrooms (Braswell et al, 2001).

Becker and Hativa (1994) estimated that half the current dollar investments for software in the United States are used to provide ILSs. Among elementary schools in the United States with networked computer systems, two-thirds are using ILSs. It is estimated that 20% of all elementary schools, a total of 10,000 schools, have ILSs installed.

Schools are employing ILS technology in the hope that CAI will improve student achievement. Van Horn (1997) exclaimed, “Like it or not, parents, administrators, and school board members firmly believe that technology needs to crank up sagging achievement test scores in order to justify the substantial money spent on it” (p. 584). He further admonished, “Improving test scores is not why we have technology in schools” (p. 584). In spite of this belief, Van Horn gave pointers for using technology effectively: (a) Use it! If you don’t use it, it won’t be effective, (b) Don’t use it for trivial things such as a typing, non-academic games, and drawing packages, (c) Use it in a way that allows students to practice many skills at once, and (d) Concentrate your resources; use with a specific group of students.

Johnson and Donley (1996) claimed that if technology is well implemented it can improve instruction and achievement. They further stated that in order to positively impact achievement the following criteria must be met: (a) Students have sufficient time to use the technology, (b) Teachers are involved and well-trained, (c) Technology is integrated with regular classroom instruction, (d) Teachers have adequate technological support, and (e) Technology and software match the curriculum and teaching methods.

Robinson (1991) also cited critical success factors for the use of ILSs. The technology must be readily accessible, teachers must be committed and motivated, and on-site support must be available. Other suggestions for implementation include: (a) Use of instructional technology should correspond with the schools' goals, (b) Each classroom needs several computers, (c) Software purchases should be reviewed, and (d) Decisions should be made regarding what it is the computer is to do, whether drill and practice or simulations (Ornstein, 1992).

Gilman (1991) stated that ILSs should focus on basic skills and that several issues, including integration and staff training influence the effectiveness of an ILS. He further stated that use of ILSs in a lab tends to isolate the system from the rest of the school. Mann et al. (999) in their study of a statewide initiative in West Virginia concluded that, "Students who had access to Basic Skills/Computer Education (BS/CE) computers in their classrooms (the "distributed" pattern) did significantly better than students who were taught with BS/CE equipment in lab settings" (p. 13). Sherry (1990) concurred, stating that optimal educational benefits cannot result when the ILS is perceived as separate from the school's curriculum plan.

Research on Student Achievement and the Use of Integrated Learning Systems

Despite the optimism surrounding instructional technology and all the advice for using it to enhance achievement, "definitive data relative to the pedagogical effects of CAI as an instructional medium remain elusive" (Burns & Bozeman, 1981, p. 35). Studies can be cited that tout the effectiveness of ILSs as well as those that state that ILSs are not beneficial to student learning. Furthermore, many researchers (Becker, 1992a;

Sherry, 1990; Trotter, 1990) have found much of the research on ILS effectiveness to be so seriously flawed that the results are meaningless.

There is no shortage of anecdotal reporting of academic success with ILS implementation. Compelling examples cited by vendors and implementing districts are abundant. The Kickstart Initiative (National Telecommunications and Information Administration, 1996) provides three such examples.

In Union City, New Jersey, students at Christopher Columbus Middle School were performing at unacceptable levels. In fact, the operation of the school was about to be taken over by the state. A few years after the infusion of technology, test scores rose to above average in reading, language arts, and math. Similar results were cited for Chula Vista's Clear View Elementary School in California. Two and one half years after implementation of technology in the classroom, students at the school rose from the 9th percentile to the 80th percentile.

Rosa Parks School was at the point of closing. Students' scores were among the lowest in the archdiocese. Over a three-year period scores rose to above average in this now high-achieving parochial district. The change is credited to an effort to improve curriculum, integrate technology, and expand the school's communication structure.

While these examples appear powerful, they should be examined with skepticism. Bentley (1991) urged caution because, "Most of the promises that ILS advocates have made about system capabilities are unsubstantiated" (p. 25). He further stated that the majority of research in this area is tainted by conflict of interest because it is conducted by the companies that are marketing the products or by districts that have invested substantially with the hope of improving student achievement. While these anecdotal reports are compelling examples of success, they must be viewed merely as a starting point for true research.

Potter and Small (1998) shares a case study in which a school followed achievement test scores for two years after implementation of a Computer Curriculum Corporation© reading program. The outcome criterion was percentage of students meeting the minimum standard on the Stanford Achievement Test. Prior to implementation 65.5% of students met this goal; two years after implementation 74.4%

were successful. Chapter One, the federal program, has a minimum requirement of an NCE gain of one year during a school year. After two years of implementation this school reported a gain of 6.65 years in total reading and 6.57 years in reading comprehension. These were the best scores ever recorded, and they followed the implementation of a 100-minute language arts block coupled with a 20-minute computer period.

In a report on progress achieved by magnet schools in Wake County, North Carolina, toward reducing the achievement gap between minority and non-minority students, Baenen (1995) stated that use of technology was a major instructional strategy. End-of-grade tests were used to measure achievement. Three of six elementary schools showed reading achievement at or above grade-level for both minority and non-minority students. Five of six schools had students performing at or above grade-level in mathematics. The gap between minority and non-minority students was reduced in four of six elementary schools for reading and in three of six for mathematics. Baenen reported that a variety of technology components were being implemented and concluded that achievement was generally improving.

In a study of effectiveness of an ILS, Bender (1991) reported that while there was no statistical significance between the control and experimental groups, all but one group did better than expected on the posttest by as many as 20 standard scale points, and students in the bottom group had an average gain of more than 100 standard scale points. When studying reading ability, students entering fourth, fifth, and sixth grades benefited more from instruction supplemented by problem-solving and simulation software than from instruction supplemented by more traditional methods.

Sixty-five independent evaluations have shown the average effect of tutoring programs in mathematics is 0.40, or four month's gain, while the median gain is substantially lower. Still, "Compared to peer tutoring, adult tutoring, increasing the length of the school day, and decreasing class size, an average CAI program produces the greatest gains per $100.00 of instructional expense" (Niemiec et al., 1986, p. 751).

Mann and Shafer (1997) report an extensive study of 55 New York State School Districts. Data were compiled and analyzed from 4,401 teachers, 1,722 students, 159 principals, and 41 superintendents. Performance on a college-preparatory examination was the dependent variable, while use of instructional technology was the independent variable. Forty two percent of the variation in math scores and 12% of the variation in English scores could be explained by the addition of technology. The most significant gains among elementary students were in sixth-grade math. There was a strong correlation between increased use of technology and higher scores on the state's comprehensive assessment. The study was not experimental; there was no control group. Neither were there performance measures from students before they ever saw computers. However, all the data (quantitative, qualitative, and longitudinal – plus anecdotal reports) pointed to one conclusion: increased technology supports student achievement. Also reported were findings that, at elementary and middle schools, technology had a greater impact in smaller schools, and the success of the technology was strongly correlated with the teacher's initiative and enthusiasm.

In a study of ninth-grade students, Schlago-Schirm (1995) found significant gains on the New Jersey Early Warning Test from pretest to posttest when CAI was the intervention. No study was conducted, however, to determine if regular classroom instruction would be more or less effective as an intervention.

Becker (1992b) found that ILSs appear to work best for students at the extremes; either high achievers or low achievers. They appear to be much less likely to help students in the middle of the class distribution. This results in overall effectiveness being rated modestly positive. The negative effect of the middle group counterbalances the positive effects at the lower and upper ends of the distribution.

Becker (1992b) conducted a randomized study of 16 classes of second through fifth graders. All students had 30 minutes of ILS instruction three times per week; half received the instruction in mathematics while the other half had reading/language arts instruction. Outcomes were assessed using both standardized achievement tests and a researcher-constructed, curriculum-specific test. Students were divided into three ability-level groups: low (below –0.5 standard deviation (sd) units under the school/grade-level mean), medium (between –0.5 and +0.5 sd units of the school/grade-level mean), and high (greater than +0.5 sd units above the school/grade-level mean). The mean effect sizes for each group, respectively, were +0.16, -0.03, and +0.16. In another study Becker

(1992a) randomly assigned students from the bottom two thirds of each grade cohort to either a “computer” (experimental) or “traditional” (control) math class taught by the same teacher. Again, posttests of the Stanford Achievement Test and curriculum-specific tests were used to assess outcomes. Fall pretests were controls in a regression analysis of effect sizes. The mean effect size for the Stanford Achievement Test was +0.45 while for the curriculum-specific tests it was +0.49, quite significant positive gains.

Burns and Bozeman (1981) concluded that a combination of traditional instruction and CAI is at least as effective as, and frequently more effective than, a program of only traditional methods. They further stated that many studies show normal instruction supplemented by CAI to be more effective than normal instruction alone. Often, it can be demonstrated that it takes less time for students to learn through CAI than through other methods.

In a meta-analysis of research findings relative to the pedagogical effectiveness of computer-assisted mathematics instruction in elementary and secondary schools, Burns and Bozeman (1981) analyzed 40 studies with student achievement as the outcome criterion. Subject variables included grade level, student ability level, and gender. The null hypothesis stated that the mean effect size would be equal to zero, and alpha was set at the 0.01 level. Overall mean effect sizes of 0.3388 and 0.4453 were found for drill and practice and tutorial models, respectively. These were significant at the 0.01 level.

There were other significant findings with regard to this study. A mathematics computer program of drill and practice or tutorial CAI plus traditional instruction was significantly more effective in fostering student achievement than traditional methods alone. Also, computer-assisted drill and practice was significantly more effective in promoting increased student achievement at elementary and secondary levels with high achievers and disadvantaged students but was not significant in aiding students of average achievement levels. This analysis and synthesis of many studies points to “a significant enhancement of learning in instructional environments supplemented by CAI, at least in one curricular area – mathematics” (Burns & Bozeman, 1981, p. 37).

In a four-month experimental study Orabuchi (1992) examined the effects of CAI on higher-order thinking skills. The subjects of the study were 61 first graders and 70 second graders. The dependent variable was performance on a standardized achievement test. Analysis of Variance (ANOVA) was used to determine the effectiveness of interactive software programs with regard to inferences, generalizations, and math problem solving. With respect to math problem solving, the first grade experimental group significantly outperformed the control group. This suggests that early exposure is important.

Fine, Bialozor, and McLaughlin (1991) conducted a small-scale study of the effect of CAI on SAT scores. In a rural setting, eight self-selected 12th graders comprised the experimental group. These students elected to take a CAI course. The control group was randomly selected from those students choosing not to take the course. The control group had comparable grade point averages and had been enrolled in similar classes when compared to the experimental group. The GPA of individuals in both groups was 3.1 or higher. There was a large difference between SAT scores of the two groups. The experimental group's scores ranged from 940 – 1480 with a mean of 1130, while the control group's scores ranged from 710 – 1250 with a mean of 951. Fine et al. suggested:

> computer-assisted drill may be more effective than traditional coaching due to the nature of CAI, the immediate feedback and remediation opportunities provided, and the unique ability of good software programs to allow students to progress at a rapid pace through the materials (p. 401).

A six-month evaluation of Computer Curriculum Corporation's SuccessMaker® courseware, carried out by Underwood, Cavendish, Dowling, Fogelman, and Lawson (1996) demonstrated a positive relationship between computer use and achievement. Sample schools consisted of five schools, both primary and secondary, that used the courseware for math and English. A pilot sample of students constituting approximately one third of the total number of students was selected. Standardized tests of non-verbal ability, mathematics achievement, and reading achievement were used to assess achievement level. Students who were exposed to the ILS were tested along with a control group of students who had not been exposed to the ILS. Tests were administered at the start (pretest) and the conclusion (posttest) of the six-month trial period.

For both primary and secondary students there was an improvement in mathematics and reading performance over the trial period for all pupils. Analysis of covariance was calculated using three variables: pre-trial performance, chronological age, and ILS use. For primary students math scores were 9.15% higher for the ILS-using group. For secondary students the ILS-using group performed significantly better that the control group achieving an average post-trial mathematics score that was 5.44% higher. The effect size for both primary and secondary mathematics was +0.4, a substantial positive effect on student achievement. There were no significant age or gender differences in performance gains, and ILS courseware did not benefit one achievement group over another any more than standard practice.

Mann et al. (1999) examined the effectiveness of a statewide technology initiative in West Virginia, Basic Skills/Computer Education (BS/CE). They sought to answer the question, "How much value can be added on a statewide basis from a sustained technology initiative?" (p. 14). The initiative was massive. Beginning with kindergarten students in 1990-91, hardware and software were installed in schools and teacher training began. The three basic components included: (a) software (IBM or Josten's Learning) that focused on the State's basic skills, (b) enough computers to afford students easy and regular access, and (c) training for teachers in the use of software and the use of computers in general. The basic skills emphasized were in the areas of reading, mathematics, and writing.

The study of the effectiveness of this initiative was a retrospective longitudinal study. Up to five years of data were collected. Schools were used as the initial stratifier for the sample. Eighteen schools ranging from high to low in achievement and socioeconomic status, representing all four geographic areas, and with proportional use of IBM and Josten's in comparison to the State were selected. Fifth graders in these schools were studied because they had all been exposed to BS/CE instruction and three years of test data were available for investigation.

The amount of student use of the computers was determined by (a) surveys of students, (b) interviews with teachers, (c) interviews with principals, and (d) interviews with selected early grades teachers. Stanford-9 test scores were used to measure student achievement. Scaled scores were selected for study because they were nationally normed for comparison purposes and they provided the opportunity to compute gain scores.

The model for the study was, “Software and computer availability and use + attitudes toward computers + teacher training and involvement in technology implementation decisions = predicted change in achievement test scores” (p. 24). The data indicated that the more of each of the model components that the student experienced, the higher the gain score on the Stanford-9. The regression model accounts for 11% of the total variance in the basic skills on the Stanford-9 Achievement Test.

Mann et al. (1999) concluded that the data indicated that all children improved academically as a result of BS/CE and that the children who were most needy benefited the most. Children without computers in the home made the biggest gains. Further indications were that there was no difference in gain scores between white and non-white students, and in math and reading there were no gender differences.

Wenglinsky (1998) conducted a study based on data drawn from the 1996 National Assessment of Educational Progress (NAEP) in mathematics to determine the impact of computer use on mathematics achievement. Students were divided into six types of subgroups including ethnicity, gender, and economic status. Weekly or more frequent use of the computer was considered to be true access. Use of the computer was also classified into two categories: lower-order and higher-order. Drill and practice was considered to be lower-order for both fourth and eighth graders while higher-order uses were learning games for fourth graders and simulations and applications for eighth graders.

The findings of this study were mixed. The use of computers to teach higher-order thinking skills was positively related to academic achievement and the social environment of the school. Additionally, it was shown that teachers’ professional development in technology was positively correlated to academic achievement in mathematics.

Teachers’ professional development and using computers for higher-order thinking skills were each associated with more than one third of a grade level increase for eighth graders. The effect size for eighth-grade use primarily for applications was +0.42.

At fourth grade the associated grade-level increase was one tenth; an effect size of +0.15 was associated with use of computers for learning games at fourth grade.

Computers seem to be associated with significant gains in mathematics achievement when they are used for higher-order thinking and when teachers have had sufficient training to manage their use appropriately. Wenglinsky's study (1998) indicated that technology can be effective in promoting academic achievement, depending upon how it is used.

Taylor (1999) investigated the link between time spent on an ILS and subsequent exam performance. Students aged 11 to 13 attending one school were studied. During the period of the investigation, the ILS was used by the mathematics department. Students were scheduled to spend one 35-minute period per week using the ILS, Learning Expedition. They were given opportunities to spend extra time with the ILS if they so desired. The log-on time was recorded in minutes by the program's management system. The variance in time due to optional time periods provided the opportunity to investigate the effect of time spent using the ILS and student performance on an achievement test.

The model for the study included two primary explanatory variables: the initial level of attainment and time spent on the ILS. Outcomes were measured using the National Foundation of Educational Research scores on verbal, non-verbal, and quantitative tests. Non-verbal and quantitative tests were expected to be affected by math achievement. This simple model was modified to allow for the possibility that the effect of an ILS on achievement might vary according to gender.

Multiple regression analysis was employed. The dominant explanatory variable, as expected, was pupil's prior level of attainment. Time spent on the ILS (for boys and girls combined) was positively related to the level of achievement and is significant at the 1% level or lower. For this study, which was limited, the data revealed that the effect of ILS hours on level achieved was stronger for girls than for boys.

In a study of 1179 students in grades K – 6, Gilman (1991) reported mixed results of achievement when Wasatch® and other software was used in a lab setting. The results for each grade level were as follows: (a) Grade one demonstrated significant losses in math and reading subtests, (b) Grade two experienced significant losses in almost all measures, (c) There were several significant losses and a one-year gain in reading at third grade, (d) Grade four saw significant gains in all areas except reading and science, (e) Significant gains in math, language arts, and science were demonstrated in fifth grade, and (f) Grade six experienced significant gains in math. Gilman further reported that year-to-year means of students were often lowered during the first year of implementation. He offered possible explanations for this phenomenon. He surmised that time normally devoted to basic skills instruction may have been used in teaching students to keyboard and to use the ILS. Another possible problem could have been that the ILS may have taught concepts not measured on the achievement test. A third possibility offered was that use of an ILS might have limited one-to-one teacher/student contact. Gilman concluded that the introduction of an ILS at Mount Vernon Elementary had not produced significant gains in student achievement.

The National Educational Longitudinal Study, as related by Weaver (2000), followed a cohort of 25,000 students who were eighth graders in 1988. This cohort was studied at two-year intervals. Students were given cognitive tests, and questionnaires were administered to each student, two of the student’s teachers, one of the student’s parents, and an administrator from each school. The responses to surveys using an ordinal scale from one (never) to three (often) were used to measure computer use.

The goal of the analysis was to determine the longitudinal effects of computer use in math and science classes. A battery of cognitive tests were used to determine achievement. The first and second follow-up tests were adaptive tests in order to base scores on the performance of each student in the base-year cycle. This Item Response Theory (IRT) allowed for comparison in improvement across years, both within and across grades.

There was a statistically significant correlation between higher average computer use and IRT scores in both math and science. Although the correlation coefficients were very small (ranging from 0.0239 to 0.0519) they were statistically significant due to the large sample size (13,120 students).

While there are many studies demonstrating positive results from ILS implementation, Gilman (1991) and other researchers reported mixed or negative results.

Holland (1993) stated that the “data are spotty” on effectiveness (p. 112). Ornstein (1992) made a stronger statement, “Because of far-reaching intellectual ramifications of the computer, the use of these machines for practice and drill is a waste of time, money, and potential” (p. 30).

Braswell et al. (2001) in *The Nation’s Report Card: Mathematics 2000* found few significant relationships between computer availability and students’ mathematics performance at grades 4, 8, and 12. This report is based upon National Assessment of Educational Progress (NAEP) scores.

For eighth graders, schools indicating that computers were available at all times in the classroom scored lower, on average, than schools who did not indicate this level of availability. Eighth graders whose teachers reported using computers primarily for demonstrating new math topics or for simulations and applications had higher mathematics scores, on average, than students whose teachers reported using computers primarily for drill or learning games. In addition, for fourth graders, the use of computers for drill and for games was associated with lower average scores than not using computers at all for instruction. This report did not address, however, the association between the amount of time spent using the computers and achievement, simply the availability of the computers with achievement.

VanDusen and Worthen (1994) conducted a large-scale study of six elementary schools across the United States over a two-year period. They collected achievement and attitude data from 4,612 students in reading and math. Students were randomly assigned to one of three implementation conditions: (a) good – greater than 30 minutes per week and completed more than three lessons per week with active integration, (b) weak – 10 to 30 minutes per week and completed two to three lessons per week with no active integration, and (c) control – no ILS use. One-way ANOVA were performed to determine the effects of time and lesson completion on individual achievement. It was found that the overall implementation had a slight but nonsignificant impact on math; there were no significant differences between students receiving greater than 30 minutes per week and the control group.

DeVaney (1996) examined the relationship between the use of technology and student computation and geometry achievement. The study consisted of 956 eighth graders. The independent variables were: availability of computers, type of computer use, and frequency with which computers and calculators were used. The dependent variable was the geometry subscale of the Trial State Mathematics Assessment for Mississippi. The results demonstrated significant negative associations between the variables of frequent use of computers and using computers for drill and practice and geometry achievement.

Becker (1992b) provided some insight into the negative effects of technology use. On first inspection, he stated, ILS programs provide a motivating environment with immediate feedback. Some reports contend, however, that the repetitive nature of their design quickly diminishes enthusiasm. He further contended that ILSs challenge appropriately, diagnose learning difficulties accurately, and provide immediate feedback. However, the very individualization that is so prized makes it more difficult for teachers to provide a variety of support activities. This also makes it more difficult for the teacher to establish appropriate instructional relationships because the students in his/her classroom are working on a wide range of tasks and skills.

Wenglinsky (1998) acknowledged that the present research indicates that students who use CAI demonstrate higher levels of academic achievement than students who do not have access to CAI. He also stated, however, that research on cost effectiveness indicated that gains in academic achievement were not proportionate to the costs of buying and maintaining the necessary software and hardware. In fact, it has been suggested that use of CAI may interfere with the social aspect of education thus interfering with the learning process.

Methodological Problems and Suggestions for Future Research

Examination of existing research quickly revealed conflicting reports of the effectiveness of ILS implementation. Furthermore, many of the existing studies were called into question due to a variety of methodological problems.

Becker (1992a) did a thorough job of demonstrating problems with several studies. Among the problems he cited were: inclusion of students who repeated grades, inadequately normed standardized tests with no comparison group tested during a similar interval, and little description about program implementation.

A study of World Institute for Computer Aided Training (WICAT) conducted in Chicago showed effect sizes of +0.3 to +0.4 for a single year. Becker (1992a) pointed out, however, that some schools reported only on a small number of students. Also, a couple of schools' data were reported on hundreds of students, but positive effects seemed to have resulted from low gains during prior (comparison) years. With correction for these errors, the effect size becomes –0.1 to +0.1. Another WICAT study, conducted in Texas demonstrated a positive relationship between ILS use and achievement. An oddity here was that most, if not all, of the effectiveness occurred during the first few months of the intervention.

Vendor-supplied studies by Computer Curriculum Corporation© showed astronomical effect sizes between +0.6 to +1.6. There were, however, two analytic problems. Cases that showed sharp declines were eliminated while correspondingly large gains were not excluded. There was also an exceptionally high attrition rate. Scores were reported for only 60% of the students using the program.

In Calvert County, Maryland, achievement test ranking for the district rose from 14th to 3rd in seven years, a feat the district credited to implementation of Computer Curriculum Corporation© software. The results were startling: More than 83% of students scored above the 50th percentile, and the first three stanines in reading and math had been virtually eliminated. However, upon inspection Becker (1992a) discovered that the ILS implementation years were not compared to years immediately prior to implementation but to scores obtained three years earlier. Another problem was failure to take an overall upward trend in achievement into account. Calvert County did rise faster than the state, but this differential rise was even larger prior to the implementation of ILS.

Becker (1992a) also investigated randomized experiments conducted on use of Computer Curriculum Corporation© software. The design of these experiments was truly experimental with students being assigned by chance to ILS-using or traditionally taught (non-using) groups. One of these studies, conducted by Educational Testing Service (ETS), reported effect sizes for mathematics computation accumulating to +0.7 over three years. The effect size for total mathematics was somewhat lower (although still substantially positive) because the effect size for mathematics concepts and applications was close to zero. Becker (1992a) found this study flawed because of use of a non-standard and upward-biasing standard deviation measure to calculate effect sizes. The recalculated effect sizes (+0.59 for three years in math computation and +0.27 for three years in total math) while substantially lower, were still significant.

Wenglinsky (1998) cited several methodological problems with existing research. One was that technology was treated as an undifferentiated characteristic; no distinction was made among various uses of technology. Another was that evaluations often used a poor measure of academic achievement, unvalidated tests developed for that particular study. He also cited failure to randomly assign subjects. Lastly, he referred to the sizes of the studies; most are not nation- or even statewide. Bork (1991) concurred, stating that, while many believe that technology is capable of improving education, the evidence is weak. Most comparison studies have been done with small groups and “a few hundred students is entirely inadequate. If we are to compare carefully alternate treatments in education, thousands are required to obtain even minimal effective evidence” (p. 12).

Olson and Krendl (1990) offered several suggestions for researchers studying the impact of technology on student achievement. They state that a longitudinal design would enable researchers to follow the progress of students who received CAI compared to students who did not. The inclusion of a control group and/or longitudinal tracking would provide a better indication of the effectiveness of CAI. Olson and Krendl (1990) further cautioned that assessment based on standardized test scores should be interpreted with care and that how software is used should be clarified for future research. Finally, they stated that researchers should exercise care to avoid an increase in attention given to students participating as subjects in the study and that isolating gender in experimental design may assist understanding of effective use for males and females.

Summary

This chapter began with a review of the historical development of computer-assisted instruction. Technology has evolved from electro-mechanical “teaching machines” to education within the computer industry to freestanding computers in classrooms to microcomputers in classrooms connected via local- and wide-area-networks. Teachers have seen computers move out of labs and into classrooms with the implicit expectation that the technology would become a transparent tool that would improve student achievement.

A wealth of information exists describing integrated learning systems, a variety of uses for the corresponding software, and proper implementation. Further, there is no shortage of logical reasoning explaining why ILSs should improve student achievement if properly implemented. It is apparent, as well, that school systems are buying into this reasoning with a significant investment of funds.

There is a growing body of research on the use of integrated learning systems to enhance traditional teaching methods. There are a number of studies that demonstrate the positive impact of ILS use on student achievement; however, other studies show negative effects, and thoughtful researchers have pointed out distinct flaws in much of the research. This chapter concluded with suggestions for improving the quality of research with regard to ILS use and student achievement.

As the infusion of technology becomes more and more prevalent, good studies of its use in the educational realm become more and more important. Additional information is imperative to guide the use of ILSs for maximum student benefit. It is hoped that many of the flaws pointed out in this review of literature can be avoided and that some of the suggestions for developing quality studies can be employed.

CHAPTER 3
RESEARCH METHODOLOGY

This chapter describes the methodology and procedures used in this study of mathematics achievement and the use of an integrated learning system. It is organized into the following sections: research design, population, student achievement, description of Greeneville City integrated learning system (ILS) implementation, description of courseware, data collection, and data analysis.

Research Design

Gall, Borg, and Gall (1996) identified four types of knowledge research that contribute to education: description, prediction, explanation, and improvement. This study proposes to contribute information about improvement in educational practice by analyzing the effectiveness of an intervention. It is a use-based evaluation study that will examine the effectiveness of the recently implemented ILS mathematics courseware in grades one through six at the four elementary schools and one middle school in the Greeneville City School System.

This is a quasi-experimental, correlational study, employing analysis of variance (ANOVA) to analyze the data in an effort to determine the extent of contribution of the integrated learning system to students' mathematics achievement and to students' gains in mathematics achievement. ANOVA is a statistical procedure used to compare the amount of between-groups variance in individuals' scores with the amount of within-groups variance (Gall et al, 1996). Stevens (1996) offers two benefits to factorial ANOVA. The first is that factorial analysis enables the researcher to examine the joint effect of independent variables. This information cannot be gathered by running separate one-way analyses. A second advantage of factorial designs is that they can increase power by reducing error or within cell variance. Post hoc testing will provide differentiation among variables. The Bonferonni inequality is a good procedure according to Stevens, who stated that, when the number of dependent variables is less than or equal to seven, the overall type I error rate for the set of t tests will be less than alpha. "The

Bonferroni inequality says that the overall alpha level for a set of tests is less than or equal to the sum of the alpha levels for each test." (p. 160). This procedure will provide information about both the magnitude of and statistical significance between variables. The intent is to evaluate the effectiveness of the integrated learning system as implemented and to provide information regarding maximum effectiveness for use in future decision making.

Gall et al (1996) defined educational evaluation as "the process of making judgements about the merit, value, or worth of educational programs" (p. 680). The purpose of educational program evaluation is to find out, specifically, how valuable a particular program may be (Raizen & Rossi, 1981). Isaac and Michael (1987) stressed the importance of this type of research "to make rational choices between alternative practices, to validate educational improvements, and to build a stable foundation of effective practices as a safeguard against faddish but inferior innovations"(Foreword). The basic question is whether a program produces more of an intended effect, in this case math achievement, than would have occurred without the program.

Gall et al. (1996) classified the type of evaluation conducted here as "goal-free evaluation" because the research is conducted to investigate the effects of a program without regard to the program's stated or proposed goals. Gall et al. further stated that, "If the evaluation study is done to answer questions primarily of interest to you [the researcher], you will need only to clarify for yourself why the study is being done" (p. 683).

With regard to the objectivity of the researcher, Raizen and Rossi (1981) stated that, while it is obviously desirable that the researcher be someone who is not deeply committed to or involved with the program being evaluated, neither should he or she be so far removed from the program as to be dispassionate to the point of disregard for the objectives and values of the program. They conclude that "good evaluators must balance precariously between an intimate and responsible knowledge of the program and a distance from it that will permit them to see its strengths and weaknesses" (p.36). While I was not responsible for selecting or implementing the Computer Curriculum Corporation© (CCC) program, as an administrator in the Greeneville City School

System, I am knowledgeable about the program, its stated capabilities, perceived capabilities, and use as an instructional tool. While the results of this study will likely prove useful to the administration of Greeneville City Schools, the decision to evaluate the implementation of an integrated learning system was mine, initiated by questions that occurred to me while compiling data during an internship project for the Greeneville City System.

Population

The Greeneville City School System serves the students of Greeneville, Tennessee, a small town located at the foothills of the Unaka Mountains in Northeastern Tennessee. During the time period of the data gathered for this study, 1997-98, 1998-99, and 1999-2000, an average of 2566 students attended the four elementary schools, one middle school, and one high school. Of this number, 575 were second, third, or fourth graders in 1997-98. Students who took the *Terra Nova* all three years of the study were included in the final analysis, resulting in a population of 348 students. Although all scores and times were analyzed for each student, classifying this group as a population, the group was also considered to be representative of future Greeneville City students. The population breakdown is shown in Table 1.

Table 1

Population Demographics of Elementary Schools

School	Average Attendance	Male (%)	Female (%)	Minority (%)
School #1	347	49.4	50.8	3.7
School #2	385	53.1	46.9	13.4
School #3	197	51.5	48.5	20.8
School #4	339	49.7	50.3	5.5

Students were tracked longitudinally; therefore, data were collected on students who were second-, third-, and fourth-graders attending one of the four elementary schools during the 1997-98 school year, and these students were followed as third-, fourth-, and fifth-graders the next year and as fourth-, fifth-, and sixth-graders during the 1999-2000 school year. Students who did not participate in the CCC program for the last two years were excluded, as were those who were absent during administration of *Terra Nova* for any of the three years. Students who were retained for the 1998-99 or 1999-2000 school years were also excluded. The population consisted of 348 students who participated in CCC mathematics instruction and who completed the Terra Nova in 1997-98, 1998-99, and 1999-2000.

Classroom settings for students in grades one through five varied. Some classes were self-contained, single-grade classrooms; others were self-contained, multi-age classrooms, while in others students had different teachers for instruction in various disciplines. All students in grades one through five received a minimum of one hour of daily mathematics instruction.

Sixth-graders received a minimum of 50 minutes of math instruction per day. Modified grouping was used with teachers providing curriculum and instruction appropriate for the student's ability level. Additionally, students with identified needs were enrolled in math investigations, a daily 40-minute course designed to provide supplemental instruction for the student's need area.

Student Achievement

Student achievement was measured using the *Terra Nova*, CTBS-5 edition, Forms I, J, and K published by McGraw Hill (R. Wankel, personal communication, February 20, 2002). This is a nationally-normed achievement test adopted by the State of Tennessee for purposes of measuring student achievement and gain (or value-added). CTB-McGraw-Hill in their *Technical Bulletin* (1997) claim a high degree of content, criterion, and constructive validity. They state, in fact

> *TerraNova* was designed and developed to provide achievement scores that are valid for most types of educational decision making. The primary inferences from the test results include measurement of the achievement of individual students relative to a current nationwide normative group and relative program

effectiveness based on the results of groups of students. Progress can be tracked over years and grades. The results can be used in a criterion-referenced manner to analyze the strengths and weaknesses of a student's achievement in each content area, to plan for further instruction, to plan for curriculum development, and to report progress to parents. The results can also be used as one factor in making administrative decisions about program effectiveness, class grouping, needs assessment, and placement in social programs. (p. 290)

The mathematics portion of the test consisted of 2 subtests: Mathematics and Mathematics Computation (*Technical Bulletin 1,* 1997). In accordance with recommendations by the National Council of Teachers of Mathematics, the Mathematics test placed major emphasis on a balance of skills, concepts, knowledge, and problem solving rather than on procedural and computational processes. This test, in grades 3 through 5, examined the basic concepts of number, operations, measurement, geometry, patterns, and data representation as well as estimation, probability, simple functions, and inferences from data. Higher levels of the test covered applications of the basic concepts, more sophisticated data representations, statistics, more complex functions and their graphs, and a greater range of problem-solving skills. "In general, the *Terra Nova* Mathematics test requires students to demonstrate mathematical power by knowing and applying a wide range of core concepts and procedures, and by creatively approaching nonroutine problems set in well-motivated, real-life situations." (p. 44). On the other hand, the Mathematics Computation test assessed the ability of students to perform operations on specific number types. This test addressed the fundamentals of computation, employing a series of computation exercises that cover arithmetic operations on number types appropriate to each grade level.

The composite score in mathematics, as reported as a scale score, was used to determine each student's mathematics achievement. These scores were also used to calculate math gain from year to year, and the scale scores from 1998 were used to form three achievement groups: low, average, and high. Students were administered the *Terra Nova* in April of 1998, 1999, and 2000 under strict guidelines and conditions mandated by the state and designed to assure standardization. Scores were taken from reports supplied by the State of Tennessee.

Description of Greeneville City Implementation

In October of 1995, the Greeneville City Schools began a system-wide technology initiative. There were several components to the initiative including hardware, support, software, and training. Five computer stations, printer, and large-screen television were placed in each classroom. These machines were networked via local area network, category 5, Novell fileservers. Personnel were hired to provide technical support both system-wide and within individual schools.

While selection of hardware was accomplished based on reliability, available technical support, and bid prices, software selection was a more involved process. It was determined that an educational software package would be purchased to serve the entire school system. The field was narrowed to three vendors: IBM, Jostens, and Computer Curriculum Corporation©. These vendors were provided a location to set up their systems, and teachers and the public were invited in for demonstrations of the software. These stakeholders provided input regarding their preference. Upon completion of this process, the consensus was that CCC would best provide for the needs of the students in Greeneville.

Training was offered in applications such as Word, Excel, and PowerPoint; it was required that all teachers attend, initially, two days of training on the CCC software applications pertaining to their discipline(s). CCC educational consultants taught these classes in Greeneville City School's training lab located at the Central Office. Substitutes were employed so teachers could be trained on weekdays during the school year. Follow-up training was conducted during the next two years. Presently, new teachers spend two days in training just prior to the beginning of the school year. The training entailed how to enroll students, how placement is determined, and how to generate and interpret the various reports available.

While there was no system-mandated time for usage, use of the program was strongly encouraged, and an expectation of 30 minutes per day was communicated.

Principals varied in their requirements for use. Periodically, the system-wide technology coordinator generated reports, and she and the assistant superintendent for instruction examined them for patterns of use or non-use. At the end of the school year reports were generated that indicated each student's time on each program and the amount of gain achieved.

Once teachers were familiar with the program, they enrolled their students using an I.D. number unique to that student. Ideally students would keep the same I.D. number from the time they were enrolled in kindergarten until they were seniors in high school. Unfortunately, this has not been the case, making tracking of students from year-to-year a more challenging endeavor.

Achievement test data were used to determine an appropriate entry level for each student. Each time a student interacted with the integrated learning system he or she entered his or her name and I.D. number. Following the first 10 sessions, the courseware analyzed each student's responses and determined an Initial Placement Motion (IPM) CCC grade-equivalent. CCC gains data are reported that reference this IPM. The management system thus determined the appropriate level in each strand from which student exercises were chosen.

Periodically throughout the year and at the end of the year teachers can generate reports that reflect each student's time on the system and achievement level. Areas of need can also be identified. In addition to reports on individual students, class reports, which are run for each class in a grade level, provide information for each grade level.

The Greeneville City Schools' technology initiative is now in its 6th year. Upgrades are accomplished consistently and with purpose. Goals are set each year and the program is under constant review. It is hoped that this study will shed light on an area not previously investigated.

Description of Courseware

The mathematics components of the SuccessMaker® courseware, Math Concepts and Skills (MCS) and Math Investigations (MI), were included for study in this research. Math Concepts and Skills is a collection of interactive exercises designed to develop and maintain essential mathematics foundations. Diagnostic logic is employed to assess the student's learning level on an ongoing basis and to then adjust the learning experience appropriately. More than 1,500 content objectives in 16 strands are addressed. Computation strands include addition, subtraction, multiplication, division, fractions, decimals, equations, and speed games. Application strands include number concepts, geometry, measurement, word problems, applications, problem-solving strategies, science applications, and probability and statistics. Math Concepts and Skills is designed to provide coverage of strategies, concepts, and skills needed for continuous progress and understanding in math for grades 1 through 8 (*Courseware Description,* 1996).

Math Investigations, building on the recommendations of the National Council of Teachers of Mathematics (NCTM), incorporates multimedia elements to provide a comprehensive problem-solving environment. Opportunities are provided for students in grades 5 through 8 to apply topics such as percent, ratio, estimation, proportion, and probability. The purpose of Math Investigations is to develop problem-solving strategies, higher-order thinking skills, and mathematical reasoning. Math Investigations consists of five to six investigations per grade level; the course components are connected directly to the NCTM standards (*Courseware Description,* 1996).

The courseware can be adapted for each student by the teacher using the SuccessMaker® Management System. A variety of reports is available including: (a) The course report which gives the student's current level, percent score for the previous session, and cumulative performance information, (b) the grouping report which gives areas of difficulty for individual students, and (c) the gains report which shows the gain over time for individual students and a mean gain for an entire group of students (*Courseware Description*, 1996).

Use of an integrated learning system is intended to provide each student with an individually prescribed educational program. Wilson (1990) stated that systems may be designed for remediation, comprehensive instruction, or for development of higher-order thinking skills. Computer Curriculum Corporation© provides all of these, and its "management system monitors virtually every student keypress and adjusts the content of the material presented based on the student's mastery of objectives."(*Courseware Description*, 1996, p. 27).

The CCC management system provides a variety of reports. The most useful and most widely used are the Course Report, Class Summary Report, and Gains Report. The Course Report was the report used to collect information for this study. Information reported for each student included time on the courseware, current average expressed as a CCC grade equivalent, and current functioning levels for each of the sixteen strands in CCC grade equivalents and percentages (*Courseware Description*, 1996).

Data Collection

Data were collected on students who were in grades 2, 3, and 4 during the 1997-98 school year. These students scores were retrieved for the 1997-98, 1998-99, and 1999-2000 school years. Students were eliminated from the study if they were retained, if they were not enrolled in CCC during the 1998-99 or 1999-2000 school years, or if they did not take Terra Nova during the time the study was conducted.

Demographic data including grade-level, school, sex, and race were collected for each student. The amount of time each student spent on Math Investigations and Math Concepts and Skills was retrieved from the Course Reports. A unique I.D. number was assigned to track students for the purposes of this study. Student achievement data were obtained from the Terra Nova reports. The students' initial scale scores, those from 1997-98, were used to categorize students as low-, medium-, or high-achieving. The criteria used by the State of Tennessee were used to determine these classifications.

Data were entered into a PC using Word 2000 © as the word processing program. These data were then transferred to the SAS © statistical package. A data file was created and various statistical procedures were applied.

Data Analysis

Scale scores from the 1998 administration of Terra Nova were used to categorize students as *low-*, *average-*, or *high-achieving*. For the 2nd, 3rd, and 4th grade cohorts, low-achieving students were identified as those students in the study with math composite scale scores below 568, 604, and 632, respectively. Those students scoring between 568 and 603 were classified as average-achieving for the 2nd grade cohort. For the 3rd grade cohort average achievement was considered to be between 604 and 635, and for the 4th grade cohort it was between 632 and 665. Students scoring above 603, 635, and 665 were classified as high-achieving for the 2nd, 3rd, and 4th grade cohorts, respectively. All students were coded by number to protect their identities.

MCS use for 1997-98, 1998-99, and 1999-2000 was categorized as low, average, or high based upon the mean use of the 1997-98 school year, 15.05 hours. Those students who had times falling below one half standard deviation of the mean, use of less than 9.64 hours, were classified as *low-use*. Those students falling within one standard deviation of the mean, use of 9.64 to 20.46 hours, were designated as *average-use,* and *high-use* was considered more than one half standard deviation above the mean, or use of more than 20.46 hours in one school year. Math Investigations was not used until 1999-2000 and then was used very little. For this reason MI use was categorized as *used* or *not used*. The cumulative time variables (1997-98 plus 1998-99 and 1997-98 through 1999-2000) were categorized in similar fashion using the means for those year spans.

Dependent variables were math composite achievement scale scores for 1998, 1999, and 2000 and gain scores. *Gain 1* was calculated by subtracting the 1998 scale score from the 1999 scale score; *gain 2* was determined by subtracting the 1999 scale score from the 2000 scale score. These gains were determined to be below, at, or above the state's expected gain in mathematics for 4th, 5th, and 6th graders that are 25, 20, and 18 respectively. There was no expectation for 3rd grade gain because testing of 2nd grade students is not mandated in Tennessee.

Descriptive statistics were generated for each group of students for time on CCC mathematics courseware and for achievement in mathematics. Data were examined for the entire three-year span, as well as for each individual year.

An analysis of variance (ANOVA) test of the dependent variables was conducted to determine significance among groups. The five models tested are shown in Tables 2 through 6.

Table 2

Model 1

Independent Variables	Dependent Variable
MCS Time (low, average, high)	Math Composite Scale Score
Gender (male, female)	

Table 3

Model 2

Independent Variables	Dependent Variable
MCS Cumulative Time (low, average, high)	Math Composite Scale Score
Gender (male, female)	

Table 4

Model 3

Independent Variables	Dependent Variable
MI Use (used, not used)	Math Composite Scale Score
Gender (male, female)	

Table 5

Model 4

Independent Variables	Dependent Variable
MCS Time (low, average, high) Gender (male, female) Ability Level (low, average, high)	Math Gain

Table 6

Model 5

Independent Variables	Dependent Variable
MI Use (used, not used) Gender (male, female) Ability Level (low, average, high)	Math Gain

Each of these models was applied to each of the three cohort groups, 2nd grade, 3rd grade, and 4th grade. Model 1 was run for all three years, 1997-98, 1998-99, and 1999-00. Model 2 was run using cumulative time from 1997-98 through 1998-99 and from 1997-98 through 1999-00. Model 3 was run for 1999-00. Model 4 was run for 1998-99 (Gain 1) and for 1999-00 (Gain 2). Model 5 was run for 1999-00 (Gain 2). Altogether 27 variations of the models were run. Bonferroni was used for post hoc testing to determine which means differed significantly.

CHAPTER 4
RESULTS

The purpose of this study was to investigate the relationship between use of an integrated learning system (ILS) as a supplement to regular mathematics instruction and mathematics achievement as measured by the math composite score and corresponding math gain on the standardized achievement test, *Terra Nova*. Data were collected from the 1997-98, 1998-99, and 1999-00 school years to examine intervention effects for boys and girls and for students of differing achievement levels.

Of the 575 students who were second, third, and fourth graders during the 1997-98 school year, 227 students were excluded from the study because they did not take the *Terra Nova* all three years of the study. The resulting population numbered 348. During the 1997-98 and 1998-99 school years all students attended one of the four elementary schools in the Greeneville City School System. During the 1999-00 school year the oldest group of students were sixth graders at Greeneville Middle School.

Two models were used to examine the effect of the use of an integrated learning system on student achievement in math. The first model explored the impact of ILS use on math composite scale scores. The second model examined the effect of use of the ILS on math gain. Math gain is the change in the student's math scale score from one year to the next. Both models took into account the gender of the student while the second model also included the students' ability levels.

The first model was run for each of the three cohort groups, each year for Math Concepts and Skills (MCS) use, cumulative time on MCS (two years and three years), and Math Investigations (MI) use for the last year, for a total of 18 variations of model 1. The second model was run for each of the three cohort groups for MCS use during 1998-99 and 1999-00 and for MI use for 1999-00, for a total of nine variations of model 2. Altogether 27 variations of two models were run.

This chapter is organized into 5 sections, each of which is associated with one or more of the guiding research questions presented in Chapter 1. Student achievement is discussed in the first. The second section includes patterns of ILS use while sections three and four are presentations of the impact of ILS use on math composite achievement and gain scores, respectively. The final section is a summary of findings.

Student Achievement

Research Question #1

How did the population's achievement scores and gain scores compare with those of the nation?

The math scale scores for the population were compared to the national norm group. These results are presented in Table 7.

Table 7

Comparison of Population Scale Scores to the National Norm Average, By Year and Cohort

Cohort Group	Mean Scale Scores					
	1998		1999		2000	
	Population	National Average	Population	National Average	Population	National Average
Second Grade Cohort	586	556	627	599	654	623
Third Grade Cohort	619	599	645	623	671	643
Fourth Grade Cohort	648	623	663	643	693	661

As shown in Table 7, the achievement of this population compared favorably with national average performance. Students in this population scored 20 to 31 scale score points higher than the national average. Math gain scores for the population were compared to the national norm group. These results are presented in Table 8.

Table 8

Comparison of Population Gain Scores to the National Expected Gain, By Year and Cohort

Cohort Group	Mean Gain Scores			
	1997-98 – 1998-99		1998-99 - 1999-00	
	Population Gain	Expected Gain	Population Gain	Expected Gain
Second Grade Cohort	42	NA	27	25
Third Grade Cohort	25	25	26	20
Fourth Grade Cohort	15	20	30	18

As shown in Table 8, the population had higher 1998-99 to 1999-2000 gain scores than the nationally expected gains. This pattern of difference was not demonstrated for the 1997-98 to 1998-99 gain scores. Among the second grade cohort, gains from third to fourth grade were slightly above expected gains between 1998-99 and 1999-2000 (2 points). The third grade cohort exceeded the expected gain between 1998-99 and 1999-2000 by six points, while the fourth grade cohort had gains 12 points higher than expected between 1998-99 and 1999-2000 (from 5th to 6th grade). Overall, this population of students performed at higher than national average levels and achieved greater than expected gains between 1998-99 and 1999-2000.

Integrated Learning System Use

Research Question #2

How much time did students spend using Math Concepts and Skills?

The amount of time each student used Math Concepts and Skills was recorded by Computer Curriculum Corporation's management system and reported on the Course Report. These times, reported as hours, were rounded to the nearest quarter hour for the purpose of this investigation. The mean times for Math Concepts and Skills are shown in Table 9.

Table 9

Mean Number of Hours Spent on Math Concepts and Skills (MCS), by Year and Cohort

Cohort Group	MCS Time Individual Years					
	1997-98		1998-99		1999-00	
	M	*SD*	*M*	*SD*	*M*	*SD*
Second Grade Cohort (*N*=113)	12.09	10.77	20.44	9.11	12.92	10.66
Third Grade Cohort (*N*=128)	20.75	11.28	11.71	5.29	10.40	5.74
Fourth Grade Cohort (*N*=107)	11.40	6.73	12.36	7.33	6.16	6.18
Total (*N*=348)	15.05	10.81	14.76	8.32	9.92	8.24

As shown in Table 9, the total MCS use for 1997-98 was higher than the following two years. While use in year 2 was very similar to year 1, (*M*=14.76 and *M*=15.05 hours, respectively), use in year 3 was considerably lower (*M*=9.92 hours). This pattern varied considerably, however, by cohort group. For example, the second grade cohort spent more hours during 1998-99 (*M*=20.44) than in 1997-98 (*M*=12.09).

<u>Research Question #3</u>

Has change occurred in cohorts' use of the ILS over the three year period?

In order to answer this research question, it was necessary to first classify students into groups based on their level of ILS use in 1997-98. Students were categorized into groups according to their use of the ILS, based on the 1997-98 mean of 15.05 hours of use. Students falling within one standard deviation (*SD*=10.81) of the mean were categorized *average-use*, while those one half standard deviation above were categorized *high-use*, and those one half standard deviation below were categorized *low-use*. Use of less than 9.64 hours per school year was considered low; 9.64 hours to 20.46 hours was average-use, while high-use was more than 20.46 hours in one school year.

Time on MCS was then used to determine whether certain groups were targeted for use or were given additional time. These results are presented in Table 10.

Table 10

Time Spent by Cohort Members Using MCS Over Time

Cohort Group	Grade	Year	MCS Time (N/%)		
			Low	Average	High
Second Grade Cohort	2nd	1997-98	59/52%	30/27%	24/21%
	3rd	1998-99	15/13%	49/43%	49/43%
	4th	1999-00	51/45%	33/29%	29/26%
Third Grade Cohort	3rd	1997-98	19/15%	55/43%	54/42%
	4th	1998-99	49/38%	74/58%	5/4%
	5th	1999-00	62/48%	61/48%	5/4%
Fourth Grade Cohort	4th	1997-98	51/48%	41/38%	15/14%
	5th	1998-99	50/47%	32/30%	25/23%
	6th	1999-00	86/80%	17/16%	4/4%

As shown in Table 10, most third graders were classified as average- or high-use with only 15% classified low-use in 1997-98 and 13% in 1998-99. Use was not this high for fourth graders. In 1997-98 48% of fourth graders were classified low-use, and subsequent years show similar patterns with 38%of fourth graders classified low-use in 1998-99 and 45% in 1999-00. Most fourth grade students (58%) were classified average-use in 1998-99 while the highest percentage of high-use for fourth graders (26%) was realized in 1999-00. Fifth grade use was observed in 1998-99 and 1999-00. Slightly more than half of fifth graders used the ILS for average or high amounts of time during these two school years. Use of MCS was quite low at sixth grade with more than 80% of students falling into the low-use category. With the exception of second and sixth graders, more than half the students were classified average- or high-use each year, with the highest use by third graders in 1997-98 and 1998-99.

Research Question #4

To what extent do boys and girls use MCS?

Table 11 shows the number and percentage of students in each time category by gender.

Table 11

Time Spent by Cohort Members Using MCS Over Time, by Gender

Grade	Year	Cohort	Gender	MCS Time (N/%)		
				Low	Average	High
2^{nd}	1997-98	Second Grade	Male	25/50%	12/25%	12/25%
			Female	34/53%	18/28%	12/19%
3^{rd}	1997-98	Third Grade	Male	6/11%	26/46%	24/43%
			Female	13/18%	29/40%	30/42%
3^{rd}	1998-99	Second Grade	Male	7/14%	21/43%	21/43%
			Female	8/12%	28/44%	28/44%
4^{th}	1997-98	Fourth Grade	Male	22/47%	19/40%	6/13%
			Female	29/48%	22/37%	9/15%
4^{th}	1998-99	Third Grade	Male	20/36%	33/59%	3/5%
			Female	29/40%	41/57%	2/3%
4^{th}	1999-00	Second Grade	Male	23/47%	12/24%	14/29%
			Female	28/44%	21/33%	15/23%
5^{th}	1998-99	Fourth Grade	Male	26/55%	14/30%	7/15%
			Female	24/40%	18/30%	18/30%
5^{th}	1999-00	Third Grade	Male	26/46%	28/50%	2/4%
			Female	36/50%	33/46%	3/4%
6^{th}	1999-00	Fourth Grade	Male	39/83%	7/15%	1/2%
			Female	47/78%	10/17%	3/5%

As shown in Table 11, MCS use was relatively even for males and females. The only exception to this pattern was fifth graders in 1998-99 where 30% of females were classified high-use while only 15% of males were so categorized.

Research Question #5

To what extent do students in different ability groups use MCS?

In order to answer this research question, it was necessary to first classify students into ability groups. Ability level group was determined using scale scores from the initial test year, 1997-98. A mean scale score and standard deviation were calculated for each of the three cohort groups. Those students falling more than one half standard deviation below the mean were categorized as the *low-ability* group, those more than one half standard deviation above the mean were categorized as the *high-ability* group, and students falling from one half standard deviation below the mean to one half standard deviation above the mean were categorized as the *average-ability* group. The classifications are presented in Table 12.

Table 12

Categorization of Cohort Groups by Ability Level Based on Mean Scale Scores from 1997-98

Cohort Group	*M/SD*	Ability Group		
		Low	Average	High
Second Grade Cohort	585.65/35.50	<567.97	567.97 - 603.33	>603.33
Third Grade Cohort	619.41/31.69	<603.56	603.56 - 635.26	>635.26
Fourth Grade Cohort	648.34/33.60	<631.54	631.54 - 665.14	>665.14

Once the ability groups were established, they were cross-classified with usage groups to address Research Question #5. Table 13 shows the breakdown of MCS use by ability level.

Table 13

Time Spent by Cohort Members Using MCS Over Time, by Ability Level

Grade	Year	Cohort	Ability Group	MCS Time (N/%)		
				Low	Average	High
			Low	6/19%	16/52%	9/29%
3rd	1998-99	Second Grade	Average	9/19%	22/47%	16/34%
			High	0/0%	11/31%	24/69%
			Low	8/29%	17/61%	3/11%
4th	1998-99	Third Grade	Average	26/38%	40/59%	2/3%
			High	15/47%	17/53%	0/0%
			Low	11/35%	9/30%	11/35%
4th	1999-00	Second Grade	Average	29/62%	9/19%	9/19%
			High	11/31%	15/43%	9/26%
			Low	18/58%	9/29%	4/13%
5th	1998-99	Fourth Grade	Average	17/35%	17/35%	14/30%
			High	15/54%	6/21%	7/25%
			Low	12/43%	12/43%	4/14%
5th	1999-00	Third Grade	Average	33/49%	34/50%	1/1%
			High	17/53%	15/47%	0/0%
			Low	16/52%	12/39%	3/9%
6th	1999-00	Fourth Grade	Average	43/90%	4/8%	1/2%
			High	27/96%	1/4%	0/0%

As shown in Table 13, for third graders in 1998-99 use was similar for low- and average-ability students, but a much higher percentage of high-ability students were classified high-use; 69% compared with 29% of low-ability students and 34% of average-ability students.

In 1998-99 fourth graders had low percentages of all ability groups in the high category, and, generally, slightly more than half the students in all ability groups were in the average range of use. A different pattern of use was observed for fourth graders in 1999-00. Students classified as low-ability fell relatively evenly into low-, average-, and high-use categories. Students of average ability had lower use with 62% falling into the low-use category. Those students classified as high-ability were more likely to experience average use (43%).

Low achieving and high achieving fifth graders in 1998-99 were less likely to use the ILS than students of average ability with 58% and 54%, respectively, falling into the low-use category. Average achieving fifth graders fell relatively evenly into the three use categories in 1998-99. The following year there were very few fifth graders classified as high-use. Fifth graders of all ability levels fell about 50-50 into low- and average-use classifications in 1999-00.

Sixth grade use was extremely low. Low-ability students were much more likely to use the ILS with 39% in the average-use category and 10% classified as high-use. More than 90% of average and high achieving students were classified as low-use.

Research Question #6

Did cohort members exhibit a difference in cumulative use of MCS?

In order to answer this research question, it was necessary to first develop categories of *cumulative use*. For purposes of this study, cumulative use of Math Concepts and Skills was calculated for two years (1997-98 – 1998-99) and three years (1997-98 – 1999-00). Cumulative use was categorized using the mean for each time span. Those students falling from one half standard deviation below the mean to one half standard deviation above the mean were classified as average use while those falling one half standard deviation above and below were classified high- and low-use, respectively. Table 14 shows this classification.

Table 14

Categorization of Cumulative MCS Use Based on Mean Cumulative Times for Two and Three Years

Years	*M*	*SD*	Cumulative MCS Time		
			Low	Average	High
1997-98 – 1998-99 (2)	29.81	15.18	<22.22	22.22 - 37.40	>37.40
1997-98 – 1999-00 (3)	39.74	19.31	<30.08	30.08 - 49.40	>49.40

The numbers and percentages of students in the cumulative use categories are shown in Table 15.

Table 15

Cumulative Time Spent by Cohort Members Using MCS Over Time for Two- and Three-year Periods

Cohort Group	Years	Cumulative MCS Time (N/%)		
		Low	Average	High
Second Grade	1997-98 - 1998-99 (2)	36/32%	41/36%	36/32%
	1997-98 - 1999-00 (3)	35/31%	33/29%	45/40%
Third Grade	1997-98 - 1998-99 (2)	29/23%	58/45%	41/32%
	1997-98 - 1999-00 (3)	32/25%	56/44%	40/31%
Fourth Grade	1997-98 - 1998-99 (2)	63/59%	19/18%	25/23%
	1997-98 - 1999-00 (3)	62/58%	36/34%	9/8%

As shown in the table, over time the second grade cohort had relatively even use in all three categories. The third grade cohort had fewer than one fourth of the students in the low-use category over time, while the fourth grade cohort had more than half of the students in low-use over time.

Research Question #7

How much time did students spend using Math Investigations?

The amount of time each student used Math Investigations was recorded by Computer Curriculum Corporation's management system and reported on the Course Report. These times, reported as hours, were rounded to the nearest quarter hour for the purpose of this investigation. The mean times for Math Concepts and Skills are shown in Table 16.

Table 16

Mean Number of Hours Spent on Math Investigations (MI), by Cohort

Cohort Group	MI Time (Hours) 1999-00	
	M	*SD*
Second Grade (*N*=113)	1.04	2.07
Third Grade (*N*=128)	0.12	0.55
Fourth Grade (*N*=107)	0.38	0.81
Total (*N*=348)	0.49	1.36

As shown in Table 16, Math Investigations was used very little during the time period of this study. In fact, it was not used at all until year 3, and then mean use was only 0.49 hours. Furthermore, 249 of the 348 subjects did not use the program at all. For this reason MI use was classified as "used" and "not used".

Research Question #8

To what extent do boys and girls use MI?

MI use by gender is shown in Table 17.

Table 17

Use versus Non-use of MI by Cohort Members, by Gender

Cohort Group	MI Use	Gender (N/%)	
		Male	Female
Second Grade	Used	21/43%	25/39%
	Not Used	28/57%	39/61%
Third Grade	Used	5/9%	5/7%
	Not Used	51/91%	67/93%
Fourth Grade	Used	22/47%	21/35%
	Not Used	25/53%	39/65%

As seen in Table 17 little difference in MI use for boys versus girls was noted.

Research Question #9

To what extent do students in different ability groups use MI?

Use by ability level is shown in Table 18.

Table 18

Use versus Non-use of MI by Cohort Members, by Ability Level

Cohort Group	Grade Level	MI Use	Ability Group (N/%)		
			Low	Average	High
Second Grade	4th	Used	11/36%	26/55%	9/26%
		Not Used	20/64%	21/45%	26/74%
Third Grade	5th	Used	4/14%	3/4%	3/9%
		Not Used	24/86%	65/96%	29/91%
Fourth Grade	6th	Used	10/32%	13/27%	20/71%
		Not Used	21/68%	35/73%	8/29%

As seen in Table 18, there were some notable differences in MI use by ability level. Of fourth grade students, those classified average ability used MI the most; 55% of these students were users compared with 36% of low-ability students and 26% of high-ability students. Fifth graders had low percentages of students using MI across all ability levels. Sixth graders showed the most marked difference in use by ability level with 71% of high-ability students using MI. Use for low-ability and average-ability students was 32% and 27%, respectively.

Impact of MCS and MI Use on Math Composite Achievement (Scale Score)

Two-way Analysis of Variance (ANOVA) was used to determine the impact of ILS use on math achievement as measured by the math composite scale score of *Terra Nova* and to determine whether gender was a factor. A two-way ANOVA was run for each cohort group (second, third, and fourth) for MCS, individual year use in 1997-98, 1998-99, and 1999-00; for cumulative MCS use, 1997-98 – 1998-99 and 1997-98 – 1999-00; and for MI use in 1999-00. Eighteen two-way ANOVAs were run to analyze the effect of ILS use on math achievement. All tables showing ANOVA results are located in the appendices. Tables of mean scores where significance was found are included in the text, but tables of mean scores where no significance was found are located in the appendices. No interaction effects were demonstrated in any of the models, so all main effects were interpreted without conducting simple main effects.

Research Question #10

Does MCS use have an effect on mathematics achievement as measured by math composite scale scores?

A comparison mean scores for second graders in 1997-98 is shown in Table 19. For second graders there was a significant main effect for time, F (2,107) = 4.077; p = .020.

Table 19

Mean Achievement (Scale Score) for Second Grade Cohort 1997-98 (Second Graders), by MCS Time Category and Gender

MCS Time Category	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Low	583.80	36.38	573.82	29.54	578.05	32.69
Average	589.42	41.00	585.22	33.58	586.90	36.10
High	607.83	48.59	597.75	19.68	602.79	36.62
Total	591.06	41.07	581.52	30.25	585.65	35.50

Bonferroni post hoc testing showed a significant difference between low- and high-use; as seen in Table 19 students categorized as high-use scored 24.74 points higher on math composite (p = .012). The coefficient of determination was .071 indicating that 7.1% of the variance in scores was accounted for by time on Math Concepts and Skills.

Mean scores for third graders in 1998-99 are shown in Table 20.

Table 20

Mean Achievement (Scale Score) for Second Grade Cohort 1998-99 (Third Graders), by MCS Time Category and Gender

MCS Time Category	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Low	610.43	24.78	619.63	15.88	615.33	20.30
Average	625.81	26.92	619.68	27.60	622.31	27.20
High	639.76	21.97	633.21	20.60	636.02	21.22
Total	629.59	26.17	625.59	24.13	627.33	25.00

For third graders in 1997-98 there was no significant effect of time on achievement, F (2, 122) = 1.924; p = .150. A significant effect was demonstrated, however, for third graders in 1998-99, F (2, 107) = 6.324; p = .003. Bonferroni testing showed significant differences between low and high (p = .013) and between average and high (p = .017). As shown in Table 20 students classified as high-use scored 20.69 points higher than low users and 13.71 points higher than average users. The Eta^2 of .106 demonstrates that 10.6% of the effect is attributable to the ILS.

Mean scores for fourth graders in 1997-98 are shown in Table 21.

Table 21

Mean Achievement (Scale Score) for Fourth Grade Cohort 1997-98 (Fourth Graders), by MCS Time Category and Gender

MCS Time Category	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Low	650.95	21.92	629.07	26.95	638.51	26.99
Average	659.47	34.70	646.09	40.60	652.29	38.12
High	672.17	30.92	670.11	29.88	670.93	29.20
Total	657.11	29.06	641.47	35.51	648.34	33.60

A main effect for time was demonstrated for fourth graders in 1997-98, F (2, 101) = 5.979; p = .004. Bonferonni testing showed a significant difference between students classified as low use and those classified as high use (p = .002). As seen in Table 21, high users outscored low users by 32.42 points. Eta^2 of .106 indicated that 10.6% of the variance in scores was due to use of Math Concepts and Skills

A comparison of mean scores for fourth graders in 1999-00 is shown in Table 22.

Table 22

Mean Achievement (Scale Score) for Second Grade Cohort 1999-00 (Fourth Graders), by MCS Time Category and Gender

MCS Time Category	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Low	667.13	36.63	662.79	34.75	664.75	35.32
Average	660.42	29.83	647.57	32.27	652.24	31.56
High	643.29	26.52	632.20	29.47	637.55	28.15
Total	658.67	33.37	650.62	34.51	654.12	34.10

Although a downward trend in achievement from low use to high use was observed for fourth graders in 1998-99, no significant effects were found, F (2, 122) = 2.942; p = .056. Significant main effects were found for fourth graders in 1999-00, F (2, 107) = 6.407; p = .002. Bonferroni testing revealed that low users scored significantly higher than high users (p = .002). As shown in Table 22, low users outperformed high users by 27.19 points. The coefficient of determination was .107 indicating that 10.7% of the variance was due to use of Math Concepts and Skills. There was no significant effect for fifth graders in 1998-99, F (2, 101) = 0.943; p = .393. Although a downward trend in achievement form low use to high use was observed for fifth graders in 1999-00, this difference was not significant, F (2, 122) = 1.992; p = .141.

Mean scores for sixth graders in 1999-00 are shown in Table 23.

Table 23

Mean Achievement (Scale Score) for Fourth Grade Cohort 1999-00 (Sixth Graders), by MCS Time Category and Gender

MCS Time Category	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Low	704.28	27.10	695.19	33.37	699.31	30.85
Average	680.00	25.45	658.40	35.81	667.29	32.93
High	696.00	*	653.67	29.54	664.25	32.09
Total	700.49	27.70	686.98	36.63	692.92	33.55

* N=1, therefore no SD calculated

A significant main effect was found for sixth graders in 1999-00, F (2, 101) = 7.359; p = .001. Bonferroni testing revealed that the decrease in achievement from low users to average users was significant (p = .001). As seen in Table 23, students classified as low-use scored 32.02 points higher on the math composite than students who were classified as average-use. Eta^2 of .127 indicates that12.7% of this variability was due to use of the ILS.

<u>Research Question #11</u>

Does two years of cumulative use of MCS have an effect on mathematics achievement as measured by math composite scale scores?

Two years of cumulative time on Math Concepts and Skills was recorded in 1999. Table 24 shows mean achievement for the second grade cohort in 1998-99 after two years of cumulative use.

Table 24

Mean Achievement (Scale Score) for Second Grade Cohort 1998-99, by Two Years Cumulative Time and Gender

	MCS Time Category					
1997-98 - 1998-99	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Low	618.28	28.25	620.67	31.28	619.47	29.40
Average	633.14	19.22	622.04	20.53	625.83	20.55
High	638.65	25.78	635.32	19.02	636.89	22.20
Total	629.59	26.17	625.59	24.13	627.33	25.00

For two years of cumulative use, a significant effect was demonstrated for the second grade cohort, F (2, 107) = 4.679; p = .011. Post hoc testing revealed that there was a significant difference between the performance of low users and high users (p = .009). As shown in Table 24, high users scored 17.42 points higher than low users on math composite. Eta^2 of .080 indicates that 8.0% of the variability was due to cumulative use of Math Concepts and Skills.

During this same time period no significant effects were demonstrated for the third grade cohort, F (2, 122) = 1.120; p = .330. There was, however, a significant effect for the fourth grade cohort, F (2, 101) = 3.632; p = .030. Table 25 shows mean achievement for the fourth grad cohort after two years of cumulative MCS use.

Table 25

Mean Achievement (Scale Score) for Fourth Grade Cohort 1998-99, by Two Years Cumulative Time and Gender

	MCS Time Category					
1997-98 - 1998-99	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Low	666.17	25.93	645.71	37.66	655.13	34.12
Average	683.78	31.63	664.60	30.38	673.68	31.67
High	665.78	24.80	679.69	36.39	674.68	32.85
Total	669.47	27.20	657.92	38.63	662.99	34.42

Bonferroni testing revealed that high users significantly outperformed low users by 19.55 points (p = .037), as seen in Table 25. The coefficient of determination, .067, demonstrates that 6.7% of the variability was due to cumulative use of Math Concepts and Skills.

Research Question #12

Does three years of cumulative use of MCS have an effect on math achievement as measured by math composite scale scores?

In 2000, where three years of cumulative use were recorded, only one cohort group experienced significant effects. No significant effects were recorded for the second grade cohort, F (2, 107) = .007; p = .993. Nor were significant effects observed for the fourth grade cohort, F (2, 101) = .330; p = .720. The third grade cohort did demonstrate significant effects from three years of MCS use, F (2, 122) = 5.199; p = .007. Bonferroni testing revealed both positive and negative effects.

Mean achievement for the third grade cohort during 1999-00 for three years of cumulative use of MCS is shown in Table 26.

Table 26

Mean Achievement (Scale Score) for Third Grade Cohort 1999-00, by Three Years Cumulative Time and Gender

	MCS Time Category					
1997-98 - 1999-00	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Low	664.79	38.73	661.06	30.42	662.69	33.76
Average	673.09	51.10	690.59	36.70	683.71	43.35
High	650.75	26.74	666.05	33.38	658.40	30.84
Total	663.04	41.17	676.39	36.50	670.55	39.02

As shown in Table 26, students in the third grade cohort classified as average users outperformed low users by 21.03 points (p = .036), but they also outperformed high users by 25.31 points (p = .004). The Eta^2 of .079 indicated the three years cumulative use of MCS was responsible for 7.9% of the variability.

<u>Research Question #13</u>

Does MI use have an effect on mathematics achievement as measured by math composite scale scores?

During the 1999-00 school year Math Investigations was used on a limited basis. Significant main effects from MI use were demonstrated for fourth graders, F (1, 109) = 6.524; p = .012. Mean scores for fourth graders in 1999-00 with regard to MI use are shown in Table 27.

Table 27

Mean Achievement (Scale Score) for Second Grade Cohort 1999-00 (Fourth Graders), by MI Use Category and Gender

MI Use Category	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Not Used	652.32	28.96	643.59	31.85	647.24	30.76
Used	667.14	37.53	661.60	36.25	664.13	36.53
Total	658.67	33.37	650.62	34.51	654.12	34.10

As shown in Table 27, users outperformed non-users by 16.8 points. The coefficient of determination was .056, indicating that 5.6% of the variance was due to MI use. Fifth graders, with an average time of 0.12 hours, saw no significant effect from MI use, F (1, 124) = .031; p = .859.

Table 28 shows mean scores for sixth graders in 1999-00 with regard to MI use.

Table 28

Mean Achievement (Scale Score) for Fourth Grade Cohort 1999-00 (Sixth Graders), by MI Use Category and Gender

MI Use Category	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Not Used	690.92	21.73	680.54	35.52	684.59	31.10
Used	711.36	30.14	698.95	36.47	705.30	33.57
Total	700.49	27.70	686.98	36.63	692.92	33.55

As seen in Table 28, sixth graders who used MI outperformed those who did not by 20.7 points, a significant effect, F (1, 103) = 9.335; p = .003. The value of Eta^2 was .083; 8.3% of the variability was due to MI use.

Research Question #14

Does MCS or MI use affect math achievement based upon gender?

For each model tested there was no interaction effect between gender and time, indicating that the use of Math Concepts and Skills and Math Investigations does not affect students differently based on gender. In only one case was gender found to be a factor; for fourth graders in 1998-99. There was a significant difference between males and females, F(1,122) = 5.164; p = .025, with females scoring 9.42 points higher than males on the math composite score.

Impact of MCS and MI Use on Mathematics Composite Gain

Three-way Analysis of Variance was used to determine the impact of MCS and MI use on academic gains in mathematics and to determine whether gender or ability level was a factor. Three way ANOVAs were run for each grade group, for MCS use in 1998-99 and 1999-00, and for MI use in 1999-00. A total of nine, three-way ANOVAs were run. There were no interaction effects for any of the models.

Although the overall sample size was large, 348 subjects, when categorizing for both gender and ability group, there were some empty cells. If no empty cells were observed, Type III Sum of Squares was used. This was the case for the second grade cohort, MCS use, 1999-00; the second grade cohort, MI use, 1999-00; the third grade cohort, MI use, 1999-00; the fourth grade cohort, MCS use, 1998-99; and the fourth grade cohort MI use, 1999-00. When empty cells were present, Type IV Sum of Squares was used. This was the case for the second grade cohort, MCS use, 1998-99; the third grade cohort, MCS use, 1998-99; the third grade cohort, MCS use, 1999-00; and the fourth grade cohort, MCS use, 1999-00.

Research Question #15

Does MCS use have an effect on mathematics achievement as measured by math gain?

Examination of gains for the second grade cohort from 1997-98 to 1998-99 showed an overall downward trend from low-use to high-use, with students classified as low users gaining, on average, 13.67 points more than high users. Further examination revealed that this was not the case for low-ability students. They demonstrated an upward trend from low-use to high-use, with low-ability, high-use students gaining, on average, 23.33 more points than low-ability, low-users. These effects were not, however, significant, F (2, 97) = 2.428; p = .094.

Table 29 displays mean gain scores for the second grade cohort from 1997-98 to 1998-99.

Table 29

Mean Mathematics Gain (1997-98 - 1998-99) for Second Grade Cohort by MCS Time Category and Ability Level

MCS Time Category	Low Ability		Average Ability		High Ability		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Low	52.67	22.76	48.89	7.88	**	**	50.40	14.97
Average	62.44	20.31	42.91	17.58	19.09	19.45	43.94	24.47
High	76.00	23.24	45.94	14.62	15.87	19.72	36.73	29.63
Total	64.48	22.50	45.09	15.07	16.89	19.41	41.67	26.16

** No Subjects

This cohort group did experience significant effects due to ability level, F (2, 97) = 51.884; p = .000. The coefficient of determination was .517 indicating that 51.7% of the variability in gain was due to ability level. Bonferonni testing demonstrated significant differences among all groups. As shown in Table 29, low-ability students gained 19.4 more points than average-ability students (p = .000). Students of average ability out gained those of high ability by 28.2 points (p = .000) during the same period.

During the time period from 1998-99 to 1999-00 gains for the second grade cohort decreased from low-use to high-use. Overall, students who were classified low-use gained 10.63 more points than those classified high-use. This trend was most notable for average-ability students. Those students classified average-ability, low-use gained, on average, 21.28 more points than average-ability, high users. These effects were not significant, F (2, 95) = 1.259; p = .289.

Mean gains for the second grade cohort from 1998-99 to 1999-00 are shown in Table 30.

Table 30

Mean Mathematics Gain (1998-99 - 1999-00) for Second Grade Cohort by MCS Time Category and Ability Level

MCS Time Category	Low Ability		Average Ability		High Ability		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Low	12.45	12.33	35.72	18.12	38.91	16.88	31.39	19.33
Average	13.67	18.56	18.44	19.98	35.67	20.64	24.97	21.77
High	18.27	26.81	14.44	11.05	30.11	23.97	20.76	22.34
Total	14.87	19.72	28.34	19.54	35.26	20.15	26.79	21.1

Ability level did prove to have significant main effects, F (2, 95) = 6.294; p = .003. Bonferroni revealed significant differences between low ability and high ability students (p = .000) and between low-ability and average-ability students (p = .010). As shown in Table 30, students classified as high-ability out gained low-ability students by 20.39 points, and average-ability students out gained those classified as low-ability by 13.47 points. The coefficient of determination was .117, indicating 11.7% of the variability was due to ability level.

An overall upward trend in gain from low-use to high-use was observed for the third grade cohort during the time period from 1997-98 to 1998-99. As a whole, those students classified high-use out gained those classified as low-use by 7.98 points. Close examination revealed varying patterns for different ability groups. Low-ability students saw an increase in gain from low-use to high-use of 6.7 points, while average-ability students in the high-use category gained, on average, 12.19 more points than average-ability students classified as low-use. Just the opposite was seen with high-ability students. Students classified average-use, gained, on average, 13.96 fewer points than low users. There were no high-ability, high-use students in the third grade cohort during this time period. These trends were not significant, F (2, 112) = .497; p = .610. Ability was not a factor for the third grade cohort for gains from 1997-98 to 1998-99, F (2,112) = .002; p = .998.

An overall upward trend in gain from low-use to high-use was also observed for the third grade cohort from 1998-99 to 1999-00. As a whole high users in this group gained 7.9 points more than students classified low-use. The gap was wider for students who were classified low-ability. Low-ability, high-use students out gained low-ability, low-use students by 13.33 points. The pattern was reversed for average-ability students. Average-ability students with average use gained 8.32 points less than average-ability, low-use students. There was only one average-ability, high-use subject. For high-ability students there was virtually no difference between gains of low-use students and high-use students. There were no high-ability students who were also classified high-use. Again, none of these effects were significant, F (2, 113) = .433; p = .650. Ability demonstrated no significant effects with regard to gain for the third grade cohort from 1998-99 to 1999-00, *F* (2, 113) = .085; *p* = .918.

For the fourth grade cohort from 1997-98 to 1998-99 little difference in gain was observed with regard to MCS use. Low-ability, high-use individuals gained 4.36 points more than low-ability, low-use individuals. This pattern was reversed for high-ability subjects. High-ability, low-use students out gained high-ability, high-use students by 6.70 points. These effects were not significant, *F* (2, 89) = 1.888; *p* = .157.

Table 31 displays mean gains for the fourth grade cohort from 1997-98 to 1998-99.

Table 31

Mean Mathematics Gain (1997-98 - 1998-99) for Fourth Grade Cohort by MCS Time Category and Ability Level

MCS Time Category	Low Ability		Average Ability		High Ability		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Low	17.39	26.35	21.47	17.36	9.13	18.37	16.30	21.47
Average	13.00	18.83	20.12	18.91	-10.67	8.91	12.34	20.61
High	21.75	8.66	18.14	23.63	2.43	29.58	14.32	24.28
Total	16.68	22.43	20.02	19.51	3.21	21.13	14.65	21.77

Ability was a significant factor with regard to gain for the fourth grade cohort from 1997-98 to 1998-99, *F* (2, 89) = 5.700; *p* = .005. The Eta^2 was .114 indicating that ability was responsible for 11.4% of the variability. Bonferroni testing revealed a significant difference between the gain of average-ability students and high-ability students. As shown in Table 31, average-ability students gained 16.81 points more than high-ability students (*p* = .004). Gains of low-ability students exceeded high-ability students by 13.46 points (*p* = .051). This was not significant at the 05 level, but it was close.

Very few subjects in the fourth grade cohort from 1998-99 to 1999-00 were classified both average-ability and high-use or high-ability and high-use. This made observation of trends impractical. For low-ability students in this group, high users outgained low users by 13.66 points. This effect was not significant at the 05 level, $F(2, 93) = .791$; $p = .457$. Gains were high for every ability level. The effect of ability was not significant, $F(2, 93) = .941$; $p = .394$.

Research Question #16

Does MI use have an effect on mathematics achievement as measured by math gain?

The effect of Math Investigations was also tested using Analysis of Variance. For the second grade cohort the gain from 1998-99 to 1999-00 for users exceeded that of non-users by 7.36 points. Examination of these gains by ability group revealed that increases in gain were seen for average- and high-ability users, 17.14 and 6.38 points, respectively. Low-ability users gained 3.75 fewer points than non-users. These effects were not significant, $F(1, 101) = 2.378$; $p = .126$.

The third grade cohort had only 10 subjects of 128 who used Math Investigations. This group was not examined for trends with regard to gain as a result of MI use.

Overall, users in the fourth grade cohort had higher gains than non-users. Low-ability users out gained low-ability non-users by 14.95 points. These effects were not significant, $F(1, 95) = .261$; $p = .611$.

Research Question #17

Does MCS use have varying affects on mathematics gain with regard to gender or ability level?

When MCS use was examined, no interaction effects were found. Therefore, it was concluded that use of the ILS, when investigating gain, did not affect boys and girls differently, nor did it have differing effects on students in various ability groups. No main effects were attributable to gender.

Research Question #18

Does MI use have varying effects on mathematics gain with regard to gender or ability level?

When MI use was examined, no interaction effects were found. Therefore, it was concluded that use of the ILS, when investigating gain, did not affect boys and girls differently, nor did it have differing effects on students in various ability groups.

Summary of Findings

To summarize the findings of this study, students in the study achieved at higher than the national average, and, overall, they accomplished greater than expected gains.

Math Concepts and Skills use was highest during the first year of implementation. The second year showed similar use, but use dropped to a lower level in year three. Highest use was experienced by third graders, while sixth graders had lowest use. There was no observable difference in use for males and females, with the exception of fifth graders in 1998-99 where 30% of females were classified high-use, while only 15% of males were in this category. Clear patterns of use with regard to ability group were not evident. Two patterns were noted; during the 1998-99 school year, high-ability third graders were more likely than average- or low-ability third graders to be in the high-use category, and low-ability sixth graders were more likely than high- or average-ability sixth graders to be in the high-use category.

Math Investigations was used very little. Use of this program began, on a very limited basis, during the 1999-00 school year. Little difference in MI use by gender was noted. There were notable differences in use of MI by different ability groups in fourth and sixth grades. Fourth graders of average ability were more likely to use Math Investigations. MI use by high-ability sixth graders was much higher than students of low- or average-ability.

Analysis of Variance was used to determine the impact of this use on student achievement as measured by the math composite scale score on *Terra Nova*. A combination of significant positive effects, significant negative effects, and no significant effects was observed. Students in the second grade cohort experienced positive effects as both second and third graders. In 1997-98, as second graders, students classified as high-use scored 24.74 points more than students classified as low-use. High-use third graders in 1998-99 outscored low users by 20.69 points, and high users outscored average users by 13.70 points. The fourth grade cohort also experienced positive effects. Fourth graders in 1997-98 who were classified high-use scored 32.42 points more than those classified low-use.

There were also significant negative effects demonstrated. In 1999-00, when students in the second grade cohort were fourth graders, low users scored 27.19 points more than high users. Sixth graders (the fourth grade cohort in 1999-00) who were classified low-use outscored those classified as high-use by 32.02 points.

No significant effects were demonstrated for the third grade cohort. These students were third graders in 1997-98, fourth graders in 1998-99, and fifth graders in 1999-00. Fifth graders in 1998-99, the fourth grade cohort, also showed no main effects on math achievement from MCS use.

Cumulative use of MCS demonstrated significant positive effects with some groups and no significant effects with others. For two years' cumulative use, the second grade cohort and the fourth grade cohort showed significant positive effects. Students in the second grade cohort who were classified high-use scored 17.42 points more than those classified low-use. High users in the fourth grade cohort outscored low users by 19.55 points. Students in the third grade cohort experienced no significant effects from two years' cumulative MCS use.

This pattern was reversed for three years' cumulative MCS use. The second and fourth grade cohorts experienced no significant effects from three years' cumulative MCS use, while significant main effects, both positive and negative, were demonstrated for the third grade cohort. Those students in the third grade cohort who were classified average-use scored 21.03 points more than those classified low-use, while average users outscored high users by 25.31 points.

For two cohort groups, second and fourth, use of Math Investigations had a significant effect on math achievement. Students in the second grade cohort who used MI scored 16.8 points more than students who did not use the program. Users in the fourth grade cohort out performed non-users by 20.7 points. Students in the third grade cohort, where use was extremely low (mean = 0.12 hours), experienced no main effect on math achievement.

Use of the integrated learning systems, Math Concepts and Skills and Math Investigations, had no significant main effects on gains in math composite scores. Significant effects of ability on math gains were demonstrated in three groups: the second grade cohort from 1997-98 to 1998-99; the second grade cohort from 1998-99 to 1999-00; and the fourth grade cohort from 1997-98 to 1998-99. No effect of ability on gain was demonstrated for the third grade cohort or for the fourth grade cohort from 1998-99 to 1999-00.

All ability groups were significantly different from one another for the second grade cohort from 1997-98 to 1998-99. Low-ability students gained 47.6 points more than high-ability students and 19.4 points more than average-ability students. Average-ability students outscored high-ability students by 28.2 points. This trend was reversed the following year. High-ability students in the second grade cohort outgained low-ability students by 20.39 points and average-ability students by 13.47 points from 1998-99 to 1999-00. For the fourth grade cohort a significant difference between average- and high-ability students' gain was demonstrated with average-ability students gaining 16.81 points more than high-ability students.

Results of this study were mixed and widely varied. There is evidence that implementation may have been inconsistent, resulting in difficulty in drawing conclusions. Careful summary and interpretation is necessary to provide meaningful conclusions and recommendations. These follow in Chapter 5.

CHAPTER 5

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

This study was conducted to explore the relationship between use of an integrated learning system as a supplement to regular mathematics instruction and mathematics achievement of second through sixth grade students including the impact of gender and ability level. A summary of the study, a summary of the findings, conclusions, and recommendations for further research and for practice follow.

Summary of the Study

Use of technology for educational purposes has increased in recent years. One means of integrating technology into the instructional program has been through use of integrated learning systems (ILSs). Many studies have been conducted which address the effectiveness of the use of computers as instructional tools. This growing body of research reveals both positive and negative effects of ILS use. Additionally, the accuracy and generalizeability of much of the research has been questioned.

The review of literature traced the history of computer-assisted instruction as delivered on ILSs. The advancement of instructional technology from its inception in the late 1950’s to the development of stand-alone microcomputers in the 1980s to the latest developments of local- and wide-area networks, CD-ROMs, and tool software were detailed. A definition of ILSs was presented along with descriptions of the technology’s potential and suggestions for successful implementation.

A review of recent research studies focused on the impact integrated learning systems had on student achievement in mathematics. Research ranged from anecdotal reporting to meta-analysis to scientific, experimental studies. Differences in student populations, hardware, courseware, and implementation plans made generalizations to current students difficult. Questions concerning research design and data analysis made the value of some findings questionable.

Although much of the current research indicates a positive impact from ILS use, the inability to generalize this information necessitates evaluation studies at the local level to make fiscally responsible, educationally sound decisions.

This use-based evaluation study investigated the use of an integrated learning system as a supplement to traditional mathematics instruction. The population was 348 students who were second, third, and fourth graders during the 1997-98 school year. Data were collected from the 1997-98, 1998-99, and 1999-2000 school years. Student achievement and gain data were measured using the standardized achievement test, *Terra Nova*. Analysis of Variance was employed to determine the effect of ILS use on mathematics achievement for students of various achievement levels, at different grade levels, and for boys as well as girls.

Summary of Findings

The findings of this study were mixed. They provide answers to the originally presented research questions, although these answers are sometimes contradictory. The following is a restatement of each research question and a summary of the findings related to it.

Research Question #1

How did the population's achievement scores and gain scores compare with those of the nation?

The achievement of this population compared favorably with national average performance. Scale scores for this population were, on average, 20 to 31 points higher than the national average, and gains for this population were generally higher than expected. Gains from 1997-98 to 1998-99 were at or slightly below expected levels, while all cohorts exhibited gains from 1998-99 to 1999-2000 that were higher than expected. Overall, this population of students outperformed the national average and achieved greater than expected gains between 1998-99 and 1999-2000.

Research Question #2

How much time did students spend using Math Concepts and Skills (MCS)?

Total MCS use for 1997-98 was higher than the following two years. Use in years 1 and 2 was similar (M = 15.05 hours and M = 14.76 hours, respectively), while use in year 3 was down considerably (M = 9.92 hours). Most grade levels averaged approximately 12 hours of use per year, with third graders experiencing higher usage (M = 20.75 hours in 1997-98 and M = 20.44 hours in 1998-99). Sixth graders had the least time of all grade levels with an average of 6.16 hours of use.

Research Question #3

Has change occurred in cohorts' use of the ILS over a three-year period?

As stated earlier, overall use was highest for third graders; only 15% were classified low-use in 1997-98, and only 13% were so classified in 1998-99. Use was lower for fourth and fifth grade students. Yet, each of these grade levels had slightly more than 50% of students in the average- or high-use categories. Use was lowest at second and sixth grades with 52% and 80% of students, respectively, classified low-use.

Research Question #4

To what extent do boys and girls use MCS?

MCS use was relatively even for males and females. The only exception to this pattern was fifth graders in 1998-99 where 30% of females were classified high-use while only 15% of males were so classified. Overall, there was no pattern of use that would indicate that gender was used to target students for MCS use.

Research Question #5

To what extent do students in different ability groups use MCS?

There were few notable patterns of usage with regard to ability level. Few third graders in 1998-99 experienced low use; third graders of low and average ability were most likely to be categorized as average-use, while high-ability third graders were most likely to fall into the high-use category (69%).

Fourth graders in 1998-99 and fifth graders in 1999-2000 (the third grade cohort) had low percentages of all students in the high-use category. There was no evident pattern of usage with regard to ability level at fourth or fifth grades.

Sixth graders had low usage overall, but sixth grade students of average and high ability level were especially unlikely to spend time on MCS (90% and 96%, respectively, classified low-use). Low-ability sixth graders were more likely than average- or high-ability students to use MCS with 52% classified low-use and 39% classified average-use. Overall, clear patterns of usage by ability level that would indicate targeting of specific groups for usage were not found.

Research Question #6

Did cohort members exhibit a difference in cumulative use of MCS?

The second grade cohort experienced relatively even use of MCS over the two- and three-year periods with about 30% of all students falling into each of the usage categories: low, average, and high. The third grade cohort had 44% and 45% of students classified as average use over the two- and three-year periods, respectively. Fewer than one fourth of the students in the cohort were classified low-use, while about 30% were classified high-use. More than half the students in the fourth grade cohort were classified low-use, resulting in their being the cohort with the lowest use over time.

Research Question #7

How much time did students spend using Math Investigations (MI)?

Math Investigations was used very little during the time period of this study. In fact, it was not used at all until 1999-2000, and then mean use was just under a half hour for the year. Furthermore, 249 of 348 subjects did not use Math Investigations at all.

Research Question #8

To what extent do boys and girls use MI?

There was little difference in MI use by gender. In all cohorts, slightly higher percentages of girls were classified not used, with the largest difference at sixth grade (the fourth grade cohort) where 53% of males did not use MI as compared to 65% of females. Overall, there was no clear pattern of usage that would indicate that gender was used to target students for MI use.

Research Question #9

To what extent do students in different ability groups use MI?

There were some notable differences in MI use by ability level. Fourth grade students classified as average-ability were more likely than fourth graders of low- or high-ability to use MI. Fifty-five percent of average-ability fourth graders used MI while only 36% of low-ability and 26% of high-ability fourth graders were users. Fifth graders had low percentages of students using MI across all ability levels. Sixth graders showed the most marked difference in use by ability level with 71% of high-ability students using MI. Use for low-ability and average-ability sixth graders was 32% and 27%, respectively. Summarizing, average-ability fourth graders and high-ability sixth graders had high MI usage.

Research Question #10

Does MCS use have an effect on mathematics achievement as measured by math composite scale scores?

Main effects attributed to MCS were inconsistent. For second graders, students classified as high-use outscored those classified as low-use by 24.74 points, and 7.1% of this variance is attributable to ILS use. Main effects were demonstrated for one group of third graders while the other had no main effects. Third graders in 1998-99 who were high users outscored low users by 20.69 points; high users also outscored average users by 13.71 points. ILS use contributed to 10.6% of this variance. Although third graders in 1997-98 who were classified high-use outscored those classified low-use by 26 points, there were no significant effects.

Each of the three fourth grade groups examined demonstrated different results of ILS use. Fourth graders in 1997-98 who were classified high-use outscored those classified low-use by 32.42 points with 10.6% of the variance attributable to ILS use. In 1998-99, fourth graders demonstrated no significant main effects as a result of ILS use. Although not statistically significant, examination of the scores shows student classified as low-use scored, on average, 61 points higher than students classified as high-use. This pattern was demonstrated again with fourth graders in 1999-2000, this time significantly. Low users outscored high users by 27.19 points; 10.7% of this variance was due to ILS use.

No main effects were demonstrated for either of the two groups of fifth graders studied. For the sixth grade group, low users outscored high users by 32.02 points with 12.7% of this variance attributable to ILS use.

When examining the effects of MCS on math composite scale scores, students at grades two and three responded more positively than students at grades four, five, and six. Second graders experienced a positive effect from MCS use; students classified as high-use outscored those classified as low-use. Of the two third grade groups examined, both demonstrated higher scores for high users than for low users. In only one group, however, were the effects statistically significant. Of the three fourth grade groups examined, one experienced a positive effect from MCS use, but the other two were either unaffected or negatively affected by use of MCS. Neither fifth grade group experienced significant effects and the sixth grade group was negatively affected with low users outscoring high users.

Research Questions #11 and #12

Does two years of cumulative use of MCS have an effect on mathematics achievement as measured by math composite scale scores?

Does three years of cumulative use of MCS have an effect on mathematics achievement as measured by math composite scale scores?

When examining cumulative MCS use, conflicting effects were once again evident. For the second grade cohort, two years of cumulative use resulted in high users outscoring low users by 17.42 points. Eight percent of this variance was due to ILS use. Three years of cumulative use, however, showed no significant effect. These results are better understood when each year's results are examined. This group of students demonstrated significant positive effects as second and third graders but a large negative effect as fourth graders.

The third grade cohort had no main effects for any of the three years studied. It is no surprise, then, that two years of cumulative use produced no effect and that the effect resulting from three years of cumulative use was conflicting with average users outperforming both low users and high users by 21.03 points and 25.31 points, respectively, with 7.9% of the variance due to ILS use.

Mixed results were also observed for the fourth grade cohort. As fourth graders, high users outperformed low users, while there was no effect for this group as fifth graders and as sixth graders low users outperformed high users. Two years of cumulative use resulted in high users outscoring low users by 19.55 points with 6.7% of this variance attributable to ILS use. Because of the conflicting effects between fourth and sixth grades, there was no main effect for the fourth grade cohort from three years of cumulative use.

Research Question #13

Does MI use have an effect on mathematics achievement as measured by math composite scale score?

When use of Math Investigations was examined users outperformed non-users in fourth and sixth grades, while fifth graders saw no main effect. It is notable that use was lowest at fifth grade with average time at 0.12 hours.

Research Question #14

Does MCS or MI use affect math achievement based upon gender?

When noting usage patterns, there was little difference between boys and girls. There was no interaction effect between gender and time when examining the impact of time on composite scale score. Therefore, the use of the ILSs, Math Concepts and Skills and Math Investigations did not affect boys and girls differently.

Research Questions #15 and #16

Does MCS use have an effect on mathematics achievement as measured by math gain?

Does MI use have an effect on mathematics achievement as measured by math gain?

When gain was examined as an indicator of math achievement, there were no conflicting results. In all cases, for both MCS and MI use, there were no significant effects for any of the grade groups studied.

There were also no significant effects on gain associated with use of Math Investigations. No examination was performed on the third grade cohort (fifth graders) due to very low numbers of users. Although not considered significant, the pattern observed for the second and fourth grade cohorts (fourth and sixth graders) was that users outgained non-users by 7.36 and 2.36 points, respectively.

Research Questions #17 and #18

Does MCS use have varying affects on mathematics gain with regard to gender or ability level?

Does MI use have varying affects on mathematics gain with regard to gender or ability level?

When noting usage patterns, there was little difference between boys and girls. There was no interaction effect between gender and time when examining the impact of time on gain. Therefore, the use of the ILSs, Math Concepts and Skills and Math Investigations did not affect boys and girls differently.

The investigation of effects of ILS use with regard to ability level had as its dependent variable gain. There were no significant main effects with regard to MCS or MI use and gain. Furthermore, there were no interaction effects found, indicating that ILS use did not have differing effects on students in various ability groups. There were two interesting albeit insignificant patterns that may bear further investigation. Low ability students appeared to benefit from MCS use. Low-ability, high-use students gained from 3.36 to 23.33 points more on math composite scale scores than low-ability, low-use students.

The second interesting pattern was observed with regard to the impact of MI use on gain. Low-ability sixth graders who were users of Math Investigations outgained non-users. Low-ability students in the sixth grade who used MI gained, on average, 14.95 points more than non-users. Fourth graders of average ability who used MI gained 17.14 points more than non-users. Users in the second grade cohort (fourth graders) who were classified high-ability gained 6.38 points more than non-users. Although these patterns were not found to be significant, they may be worthy of further investigation.

An interesting incidental finding was that gain is affected by ability level. These findings, while statistically significant, did not produce a clear pattern with lower ability students outgaining those of higher ability one year, followed the next year by a reversal in this pattern.

Conclusions

There was no clear indication that use of the integrated learning system, Math Concepts and Skills, either benefited or harmed students as measured by math composite scale score on *Terra Nova*. It is notable that negative effects associated with ILS use in this study occurred during the 1999-2000 school year, the year with the lowest overall usage. The mean time in 1999-2000 was 9.92 hours compared with 15.05 hours and 14.76 hours for 1997-98 and 1998-99, respectively. It is also noteworthy that, at sixth grade, overall usage was low (more than 80% of students were classified low-use) and that those students who did use MCS were almost exclusively low-ability students. There were 6 conclusions drawn from this study.

Conclusion #1

When examining effects of use of Math Concepts and Skills on children of differing grade levels, data demonstrate that this ILS is more effective with children in second and third grades than it is with children in grades four through six. In two of three cases, high-use students at grades two and three significantly outperformed low-use students. In the third scenario high-use students outperformed low-use students, but the difference was not found to be statistically significant. It should be noted that usage patterns may have affected results at 3rd and 6th grades. IN 1998-99 where a positive effect was demonstrated for 3rd graders, 69% of high-ability students were also high-use. The reverse was true at 6th grade. Negative effects could have been due to the fact that low-ability students used the ILS more.

Conclusion #2

While cumulative use for two years demonstrates benefit, three years of cumulative use of MCS is not beneficial. It is worth noting that, although years with negative effects counterbalance those with positive effects and two years of positive effects yield overall positive effects for the two year period, a year that demonstrated positive effects followed by a year with no main effects produced positive effects overall for the two-year period.

Conclusion #3

Clearly, use of MCS does not result in higher gain scores. However, the observed trend that low-ability, high-use students outgained low-ability, low-use students, although not statistically significant in this study, bears further examination. Enabling low-ability students to improve gains will, over time, move them from the low-ability classification into average-ability range. This is a key component of the “No Child Left Behind” legislation. (Learning First Alliance, 2002)

Conclusion #4

No interaction effects were found in any of the models run; use of the ILSs, MCS and MI, do not affect boys and girls differently, nor do they affect students of varying ability levels differently.

Conclusion #5

Math Investigations, although used on a very limited basis during the course of this study, had a positive effect overall on math composite scale scores. With the exception of fifth grade, where use was extremely low, users significantly outperformed non-users. It should be noted that fourth grade students of average ability and sixth grade students of high ability used MI more frequently than students in other ability levels.

Conclusion #6

When gain was examined, no significant effects of MI use were found, but positive trends students were observed which bear further investigation.

Recommendations for Future Study

Several recommendations for future research were prompted by this study. Controlling for usage, such as minimum usage requirements and/or use of a control group would be very desirable. Olson and Krendle (1990) state that inclusion of a control group and/or longitudinal tracking would provide a better indication of the effectiveness of computer-assisted instruction (CAI).

Although this study did provide a longitudinal view, it was a retrospective view. A true longitudinal study would be beneficial because it would provide the researcher with the opportunity to investigate a number of usage issues that were impossible to ascertain when "looking back" at the data. These include questions about the method of implementation:

1. Was implementation standardized with minimum daily requirements?
2. Were certain groups of students targeted for or excluded from use, either by design or because of convenience?
3. Did teachers plan for usage or was CAI used as "filler" when students had extra time?

A true longitudinal study would also provide the researcher the opportunity to measure attitudes toward and expectations for the ILS by administrators, teachers, and students. This study of attitudes could help the researcher develop a deeper understanding of the effectiveness. Becker and Hativa (1994) stated that recent research reveals that many teachers continue to teach the same things in the same way using computers as a diversion, or change of pace. Many technological innovations have failed because they added extra burdens to teachers' work. It would also be beneficial to know if administrator attitudes and expectations have an impact on effectiveness. Van Dusen and Worthen (1994) stated, "We believe that ILS evaluation studies may underestimate the impact of the ILS because they do not account for the level of implementation of the ILS at the particular schools they study." (p. 14).

It would have been very helpful to compare the performance of these students prior to implementation with their performance after implementation. It is, therefore, recommended that student achievement data be collected prior to the implementation of a major intervention like CAI, and tracked each year following the implementation. Focus groups of students could be identified to make this evaluation more manageable.

It would also have been beneficial to have students categorized into ability level groups using an achievement measure other than the standardized achievement test used to measure outcomes of the intervention. This would have permitted evaluation of yearly achievement as measured by scale score with regard to ability group. In this way the effect of the ILS on students of varying abilities could be evaluated by both annual scale scores and by gain from year to year.

Further study of the effectiveness of an ILS, specifically a skill-based one such as Math Concepts and Skills, is warranted. This study indicated MCS be more useful at lower grades. Further study would strengthen this finding. The effect of MCS on gain also bears further study. Overall patterns from this study, although not statistically significant, suggest that for low-ability students, the use of MCS may improve gain. There is also some evidence that higher usage increases the likelihood of positive effects; this should be investigated further.

Further study of the ILS, Math Investigations is also needed. Initial study indicates that this program has promise in boosting math achievement. A study involving more extensive usage would be quite informative. Trends were also observed that suggest that use of MI may improve gains. Further study of this topic is also warranted.

An incidental finding of this study was the association between ability level and gain. Again, results were mixed. The findings that ability level significantly affected gain could prove very informative. Further review is needed.

Recommendations for Practice

Several recommendations for practice can be made as a result of this study. The first would be that usage should be monitored for effectiveness, both by school and by classroom. Bork (1991) provided support for this recommendation in stating that, because we never test the effectiveness of the computer directly, because we always test it under specific circumstances (specific software in a specific setting) we must use extreme caution in generalizing findings. The fact that many of the findings from this study were inconsistent leads one to believe that the specific circumstances under which the ILS was implemented had an impact on the effectiveness.

Another recommendation would be to target lower grades with skill-based integrated learning systems such as Math Concepts and Skills as positive effects were seen primarily at second and third grades in this study. Also, increasing usage of Math Investigations at grades four through six with concurrent effectiveness monitoring is recommended.

There is no need to target either boys or girls with Math Concepts and Skills or Math Investigations. Neither is there statistical evidence that targeting specific ability groups would prove effective.

How can assessments be designed and carried out in technology-rich environments? Weaver (2000) stated that assessments should not be limited to comparing the effectiveness of computer-assisted instruction with conventional instruction because “these forms of assessment may fail to demonstrate that unexpected forms of cognitive change occur as a result of technology-rich environments” (p. 132). Purchase and implementation of integrated learning systems or other technological interventions are expensive endeavors and should not be employed without appropriate research and careful planning for implementation to maximize effectiveness. These programs should also be monitored and continually assessed to insure their ongoing effectiveness. Time is the most precious commodity we have in education; we must spend it wisely, employing the most efficient models to promote learning.

REFERENCES

Archer, J. (1998, October). The link to higher scores. *Education Week on the WEB.* Retrieved December 28, 2000, from http://www.edweek.org/sreports/tc98/ets-n.htm

Baenen, N. (1995). *Improving achievement through technology: Status report on the the magnet schools assistance program.* Raleigh, NC: Wake County Public Schools System. (ERIC Document Reproduction Service No. ED384342)

Becker, H.J. (1992a). Computer-based integrated learning systems in the elementary and middle grades: A critical review and synthesis of evaluation reports. *Journal of Educational Computing Research, 8,* 1-41.

Becker, H.J. (1992b). A model for improving the performance of integrated learning systems: Mixed individualized/group/whole class lessons, cooperative learning, and organizing time for teacher-led remediation of small groups. *Educational \Technology, 32* (9), 6-15.

Becker, H.J., & Hativa, N. (1994). History, theory, and research concerning integrated learning systems. *International Journal of Educational Research, 21,* 5-12.

Bender, P.V. (1991). The effectiveness of integrated computer learning systems in the elementary school. *Contemporary Education, 63,* 12-23.

Bentley, E. (1991). Integrated learning systems: The problems with the solution. C*ontemporary Education*, *63,* 24-27.

Beyond the numbers: A guide to interpreting and using the results of standardized achievement tests. (1997). Monterey, CA: CTB/McGraw-Hill.

Bork, A. (1991). Is technology-based learning effective? *Contemporary Education, 63,* 8-16.

Bracey, G. (1991). ILS research isn't helpful. *Electronic Learning, 11*(3), 16-22.

Braswell, J., Lutkus, A.D., Grigg, W.S., Santapau, S.L., Tay-Lim, B., & Johnson, M. (2001). *The Nation's Report Card: Mathematics 2000* (Report No. NCES 2001-517). Washington DC: US Department of Education. (ERIC Document Reproduction Service No. ED453086)

Burns, P.K., & Bozeman, W.C. (1981). Computer-assisted instruction and mathematics achievement: Is there a relationship? *Educational Technology, 21*(10), 32-39.

Coley, R., Cradler, J., & Engel, P.K. (1997). *Computers and classrooms: The status of technology in U.S. Schools.* Princeton, NJ: Educational Testing Service. (ERIC Document Reproduction Service No. ED412893)

Courseware Description. (1996). Sunnyvale, CA: Simon and Schuster.

DeVaney, T.A. (1996, November). *The effects of computers and calculators on computation and geometry achievement.* Paper presented at the annual meeting of the Mid-south Educational Research Association, Tuscaloosa, AL.

Division of Educational Sciences, Content, and Methods of Education, Unesco. (1987). *Glossary of educational technology terms* (2nd ed.). Lausanne, Switzerland: Presses Centrales.

Fine, L., Bialazor, R.C., & McLaughlin, T.F. (1991). An analysis of computer-assisted instruction on scholastic aptitude performance of rural high school students. *Education, 111,* 400-403.

Gall, M., Borg, W.R., & Gall, J.P. (1996). *Educational research: An introduction* (6th ed.). White Plains, NY: Longman Publishers.

Gilman, D.A. (1991). *A comprehensive study of the effects of an integrated learning system. A report prepared for the Metropolitan School District of Mount Vernon, Indiana.* Terre Haute, IN: Indiana State University. (ERIC Document Reproduction Service No. ED339363)

Holland, W.R. (1993). Is ILS worth the investment? *NASSP Bulletin, 77*(553), 112-114.

Honey, M., Culp, K.M., & Carrigg, F. (1999). *Perspectives on technology and education research: Lessons from the past and present.* Whitepaper presented at the Secretary's Conference on Educational Technology, Washington DC. Retrieved December 28, 2000, from http://www.ed.gov/Technology/TechConf/1999/Whitepapers/paper1.html

Interpreting tests (n.d.). Retrieved December 28, 2000, from http://www.quia.com

Isaac, S., & Michael, W.B. (1987). *Handbook of research and evaluation for education* *and the behavioral sciences* (2nd ed.). San Diego, CA: EdiTS.

Johnson, J.L., & Donley, J. (1996). *Magnet schools assistance program 1993-95. Evaluation report.* Raleigh, NC: Wake County Public Schools System. (ERIC Document Reproduction Service No. ED395984)

Kulik, J.A. (1994). Meta analytic studies of findings on computer-based instruction In E.L. Barker & H.F. O'Neil, Jr. (Eds.), *Technology assessment in education and training.* Hillsdale, NJ: Lawrence Erlbaum.

Learning First Alliance. (2002). *Major changes to ESEA in the no child left behind act.* Retrieved July 30, 2002, from http://www.learningfirst.org.

Mageau, T. (1991). Integrating an ILS: Two teaching models that work. *Electronic Learning, 11*(3), 16-22.

Mann, D., (1999). *Documenting the effects of instructional technology: A fly-over of policy questions.* Whitepaper presented at the Secretary's Conference on Educational Technology, Washington DC. Retrieved December 28, 2000, from http://www.ed.gov/Technology/TechConf/1999/Whitepapers/paper1.html

Mann, D., & Schafer, E.A. (1997). Technology and achievement. *The American School Board Journal, 184*(7). 22-23.

Mann, D., Shakeshaft, C., Becker, J., & Kottkamp, R. (1999). *West Virginia story: Achievement gains from a statewide comprehensive instructional technology program.* Beverly Hills, CA: Milken Family Foundation. (ERIC Document Reproduction Service No. ED429575)

McNabb, M., Hawkes, M., & Rouk, U. (1999). *Critical issues in evaluating the effectiveness of technology.* Summary of the Secretary's Conference on Educational Technology, Washington DC. Retrieved December 28, 2000, from http://www.ed.gov/Technology/TechConf/1999/Whitepapers/paper1.html

National Telecommunications and Information Administration. (1996). *Kickstart initiative: Connecting America's communities to the information superhighway.* Washington DC: Author. (ERIC Document Reproduction Service No. ED393472)

Niemiec, R.P., Blackwell, M.C., & Walberg, H.J. (1986). CAI can be doubly effective.

Phi Delta Kappan, 67, 750-751.

Olson, B., & Krendl, K.A. (1990). At risk students and microcomputers: What do we know and how do we know it? *Journal of Educational Technology Systems, 19,* 165-175.

Orabuchi, I.I. (1992). *Effects of using interactive CAI on primary grade students' higher-order thinking skills: Inference, generalizations, and math problem solving*. Denton,TX: Texas Women's University. (ERIC Document Reproduction Service No. ED390383)

Ornstein, A.C. (1992). Making effective use of computer technology. *NASSP Bulletin, 76*(542). 27-33.

Potter, L., & Small, J. (1998). Utilizing computers for reading improvement in junior high: A case study. *International Journal of Instructional Media, 25.* 383.

Raizen, S.A., & Rossi, P.H. (Eds.). (1981). *Program evaluation in education.* Washington DC: National Academy Press.

Robinson, S. (1991). Integrated learning systems: From teacher proof to teacher empowering. *Contemporary Education, 63.* 15-18.

Schlago-Schirm, C. (1995). *Does the computer-assisted remedial mathematics program at Kearny High School lead to improved scores on the N.J. Early Warning Test?.* Union, NJ: Kean College of New Jersey. (ERIC Document Reproduction Service No. ED382460)

Sherry, M. (1990). Implementing an integrated instructional system: Critical issues. *Phi Delta Kappan, 72,* 118-120.

Starr, P. (1996). Computing our way to education reform. *The American Prospect, 27,* 50-60.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Taylor, L. (1999). An integrated learning system and its effect on examination performance in mathematics. C*omputers and Education, 32,* 95-107.

Technical Bulletin 1 (1997). Monterey, CA: CTB/McGraw-Hill.

Test coordinator's handbook. (1997). Monterey, CA: CTB/McGraw-Hill.

Trotter, A. (1990). Computer learning. *The American School Board Journal, 177*(7). 12-18.

Underwood, J., Cavendish, S., Dowling, S., Fogelman, K., & Lawson, T. (1996). Are integrated learning systems effective learning support tools?. *Computers and Education, 26,* 33-40.

US Department of Education. (1997). *Investing in school technology: Strategies to meet the funding challenge.* Retrieved December 28, 2000, from http://www.edgov/pubs/techinvest/chap1.html

VanDusen, L.M., & Worthen, B.R. (1994). The impact of integrated learning system integration on student outcomes: Implications for research and evaluation. *International Journal of Educational Research, 21,* 13-24.

VanHorn, R. (1997). Power tools. *Phi Delta Kappan, 78,* 584-585.

Vargas, J.S. (1986). Instructional design flaws in computer-assisted instruction. *Phi Delta Kappan, 67,* 738-744.

Weaver, G.C., (2000). An examination of the national educational longitudinal study (NELS:88) database to probe the correlation between computer use in school and improvement in test scores. *Journal of Science Education and Technology, 9,* 121-133.

Wenglinsky, H. (1998). *Does it compute? The relationship between educational technology and student achievement in mathematics.* Princeton, NJ: Educational Testing Service. (ERIC Document Reproduction Service No. ED425191)

APPENDICES

APPENDIX A

ANOVA RESULTS TABLES, MODEL 1

Table A1

ANOVA for Second Grade Cohort, Dependent Variable: Math Composite Scale Score 1998 (Second Graders), Independent Variables: MCS Time and Gender

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	1568.853	1568.853	1.307	0.012
Time98	2	9788.33	4894.165	4.077*	0.071
Gender*Time98	2	179.918	89.959	0.075	0.001
Error	107	128456.886	1200.532		
Total	113	38899187			
Corrected Total	112	141133.54			

Note: $N = 113$

*$p < .05$

Table A2

ANOVA for Third Grade Cohort, Dependent Variable: Math Composite Scale Score 1998 Third Graders), Independent Variables: MCS Time and Gender

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	2660.422	2660.422	2.653	0.021
Time98	2	3858.12	1929.06	1.924	0.031
Gender*Time98	2	1568.799	784.4	0.782	0.013
Error	122	122324.43	1002.659		
Total	128	49236576			
Corrected Total	127	127570.875			

Note: *N* = 128

Table A3

ANOVA for Fourth Grade Cohort, Dependent Variable: Math Composite Scale Score 1998 (Fourth Graders), Independent Variables: MCS Time and Gender

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	3056.32	3056.32	3.129	0.03
Time98	2	11678.719	5839.359	5.979*	0.106
Gender*Time98	2	1202.784	601.392	0.616	0.012
Error	101	98639.094	976.625		
Total	107	45096094			
Corrected Total	106	119697.888			

Note: $N = 107$

*$p < .05$

Table A4

ANOVA for Second Grade Cohort, Dependent Variable: Math Composite Scale Score 1999 (Third Graders), Independent Variables: MCS Time and Gender

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	27.905	27.905	0.048	0
Time99	2	7284.214	3642.107	6.324*	0.106
Gender*Time99	2	780.815	390.408	0.678	0.013
Error	107	61621.458	575.901		
Total	113	44539986			
Corrected Total	112	69998.885			

Note: $N = 113$

*$p < .05$

Table A5

ANOVA for Third Grade Cohort, Dependent Variable: Math Composite Scale Score 1999 (Fourth Graders), Independent Variables: MCS Time and Gender

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	6045.875	6045.875	5.164*	0.041
Time99	2	6889.547	3444.773	2.942	0.046
Gender*Time99	2	4158.583	2079.291	1.776	0.028
Error	122	142827.524	1170.717		
Total	128	53364821			
Corrected Total	127	158761.43			

Note: $N = 128$

*$p < .05$

Table A6

ANOVA for Fourth Grade Cohort, Dependent Variable: Math Composite Scale Score 1999 (Fifth Graders), Independent Variables: MCS Time and Gender

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	2229.473	2229.473	1.944	0.019
Time99	2	2162.514	1081.257	0.943	0.018
Gender*Time99	2	2237.032	1118.516	0.976	0.019
Error	101	115806.226	1146.596		
Total	107	47158170			
Corrected Total	106	125612.991			

Note: N = 107

Table A7

ANOVA for Second Grade Cohort, Dependent Variable: Math Composite Scale Score 2000 (Fourth Graders), Independent Variables: MCS Time and Gender

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1		2295.869	2.154	0.02
Time00	2	13658.889	6829.445	6.407 *	0.107
Gender*Time00	2	410.617	205.308	0.193	0.004
Error	107	114046.64	1065.856		
Total	113	48479183			
Corrected Total	112	130269.504			

Note: *N* = 113

**p* < .05

Table A8

ANOVA for Third Grade Cohort, Dependent Variable: Math Composite Scale Score 2000 (Fifth Graders), Independent Variables: MCS Time and Gender

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	3388.025	3388.025	2.274	0.018
Time00	2	5934.564	2967.282	1.992	0.032
Gender*Time00	2	445.863	222.931	0.15	0.002
Error	122	181766.153	1489.886		
Total	128	57746438			
Corrected Total	127	19339.719			

Note: *N* = 128

Table A9

ANOVA for Fourth Grade Cohort, Dependent Variable: Math Composite Scale Score 2000 (Sixth Graders), Independent Variables: MCS Time and Gender

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	3285.357	3285.357	3.446	0.033
Time00	2	14032.863	7016.432	7.359*	0.127
Gender*Time00	2	1249.982	624.991	0.656	0.013
Error	101	96298.241	953.448		
Total	107	51493462			
Corrected Total	106	119292.243			

Note: *N* = 107

*$p < .05$

APPENDIX B

ANOVA RESULTS TABLES, MODEL 2

Table B1

ANOVA for Second Grade Cohort, Dependent Variable: Math Composite Score (Scale Score) 1999, Independent Variables: Cumulative MCS Time 1997-98 – 1998-99 and Gender

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	438.505	438.505	0.744	0.007
MCS Time1	2	5519.241	2759.62	4.679*	0.08
Gender*MCS Time1	2	836.742	418.371	0.709	0.013
Error	107	63106.276	589.778		
Total	113	44539986			
Corrected Total	112	69998.885			

Note: $N = 113$

*$p < .05$

Table B2

ANOVA for Third Grade Cohort, Dependent Variable: Math Composite Score (Scale Score) 1999, Independent Variables: Cumulative MCS Time 1997-98 – 1998-99 and Gender

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	2642.198	2642.198	2.107	0.017
MCS Time1	2	2808.457	1404.228	1.12	0.018
Gender*MCS Time1	2	256.663	128.332	0.102	0.002
Error	122	152963.605	1253.8		
Total	128	53364821			
Corrected Total	127	158761.43			

Note: N = 128

Table B3

ANOVA for Fourth Grade Cohort, Dependent Variable: Math Composite Score (Scale Score) 1999, Independent Variables: Cumulative MCS Time 1997-98 – 1998-99 and Gender

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	1476.247	1476.247	1.397	0.014
MCS Time1	2	7675.291	3837.646	3.632*	0.067
Gender*MCS Time1	2	5221.29	2610.645	2.471	0.047
Error	101	106716.145	1056.595		
Total	107	47158170			
Corrected Total	106	125612.991			

Note: $N = 107$

*$p < .05$

Table B4

ANOVA for Second Grade Cohort, Dependent Variable: Math Composite Score (Scale Score) 2000, Independent Variables: Cumulative MCS Use 1997-98 – 1999-00 and Gender

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	1753.939	1753.939	1.491	0.014
MCS Time2	2	16.086	8.043	0.007	0
Gender*MCS Time2	2	2538.769	1269.385	1.079	0.02
Error	107	125901.9	1176.653		
Total	113	48479183			
Corrected Total	112	130269.5			

Note: $N = 113$

Table B5

ANOVA for Third Grade Cohort, Dependent Variable: Math Composite Score (Scale Score) 2000, Independent Variables: Cumulative MCS Time 1997-98 – 1999-00 and Gender

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	2799.069	2799.069	2.017	0.016
MCS Time2	2	14426.76	7213.38	5.199*	0.079
Gender*MCS Time2	2	2451.389	1225.694	0.883	0.014
Error	122	169272.055	1387.476		
Total	128	57746438			
Corrected Total	127	193399.719			

Note: $N = 128$

Table B6

ANOVA for Fourth Grade Cohort, Dependent Variable: Math Composite Score (Scale Score) 2000, Independent Variables: Cumulative MCS Time 1997-98 – 1999-00 and Gender

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	1802.182	1802.182	1.602	0.016
MCS Time2	2	742.731	371.366	0.33	
Gender*MCS Time2	2	250.397	125.198	0.111	0.002
Error	101	113590.772	1124.661		
Total	107	51493462			
Corrected Total	106	119292.243			

Note: *N* = 107

APPENDIX C

ANOVA RESULTS TABLES, MODEL 3

Table C1

ANOVA for Second Grade Cohort, Dependent Variable: Math Composite Score (Scale Score) 2000, Independent Variables: MI Use and Gender

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	1367.769	1367.769	1.233	0.011
NMI00	1	7235.604	7235.604	6.524*	0.056
Gender*NMI00	1	68.258	68.258	0.062	0.001
Error	109	120894.114	1109.12		
Total	113	48479183			
Corrected Total	112	130269.504			

Note: *N* = 113

*$p < .05$

Table C2

ANOVA for Third Grade Cohort, Dependent Variable: Math Composite Score (Scale Score) 2000, Independent Variables: MI Use and Gender

Source	*df*	Type III *SS*	*MS*	*F*	Eta Squared
Gender	1	212.918	212.918	0.141	0.001
NMI00	1	47.395	47.395	0.031	0
Gender*NMI00	1	959.477	959.477	0.637	0.005
Error	122	186780.647	1506.296		
Total	128	57746438			
Corrected Total	127	193399.719			

Note: *N* = 128

Table C3

ANOVA for Fourth Grade Cohort, Dependent Variable: Math Composite Score (Scale Score) 2000, Independent Variables: MI Use and Gender

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	3273.243	3273.243	3.212	0.03
NMI00	1	9513.362	9513.362	9.335*	0.083
Gender*NMI00	1	25.957	25.957	0.025	0
Error	103	104965.576	1019.083		
Total	107	51493462			
Corrected Total	106	119292.243			

Note: $N = 107$

*$p < .05$

APPENDIX D

SAMPLE MEANS TABLES, MODEL 1

Table D1

Mean Achievement (Scale Score) for Third Grade Cohort 1997-98 (Third Graders), by MCS Time Category and Gender

MCS Time Category	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Low	592.33	71.58	617.08	32.64	609.26	47.68
Average	619.92	27.27	623.62	21.98	621.87	24.46
High	618.33	36.34	622.17	27.18	620.46	31.32
Total	616.29	37.77	621.83	26.04	619.41	31.69

Table D2

Mean Achievement (Scale Score) for Third Grade Cohort 1998-99 (Fourth Graders), by MCS Time Category and Gender

MCS Time Category	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Low	646.9	33.21	657.31	29.83	653.06	31.34
Average	639.73	37.24	642.83	34.2	641.45	35.37
High	586.33	40.46	649.5	50.2	611.6	51.44
Total	639.43	37.7	648.85	33.1	644.73	35.36

Table D3

Mean Achievement (Scale Score) for Fourth Grade Cohort 1998-99 (Fifth Graders), by MCS Time Category and Gender

MCS Time Category	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Low	672.35	27.64	650.33	43.16	661.78	37.25
Average	663.71	29.5	653.39	34.4	657.91	32.26
High	670.29	22.1	672.56	33.86	671.92	30.58
Total	669.47	27.2	657.92	38.63	662.99	34.42

Table D4

Mean Achievement (Scale Score) for Third Grade Cohort 1999-00 (Fifth Graders), by MCS Time Category and Gender

MCS Time Category	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Low	668.08	48.45	680.53	34.86	675.31	41.2
Average	661.18	30.99	673.76	39.17	667.98	35.92
High	623.5	68.59	655.67	20.03	642.8	41.08
Total	663.04	41.17	676.39	36.5	670.55	39.02

APPENDIX E

SAMPLE MEANS TABLES, MODEL 2

Table E1

Mean Achievement (Scale Score) for Third Grade Cohort 1998-99, by Two Years Cumulative Time and Gender

MCS Time Category 1997-98 - 1998-99	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Low	635.08	48.82	643.35	36.9	639.93	41.61
Average	645.88	31.66	652.85	27.98	649.84	29.56
High	633.68	38.02	647.09	37.69	640.88	37.98
Total	639.43	37.7	648.85	33.1	644.73	35.36

Table E2

Mean Achievement (Scale Score) for Third Grade Cohort 1999-00, by Three Years Cumulative Time and Gender

MCS Time Category 1997-98 – 1999-00	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Low	651.87	36.2	656.75	35.71	654.66	35.47
Average	664.81	29.38	645.29	35.14	654.76	33.47
High	658.89	35.03	649.44	33.84	653.22	34.24
Total	658.67	33.37	650.62	34.51	654.12	34.1

Table E3

Mean Achievement (Scale Score) for Fourth Grade Cohort 1999-00, by Three Years Cumulative Time and Gender

MCS Time Category 1997-98 - 1999-00	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Low	703.13	29.68	687.65	32.63	695.39	31.91
Average	694.08	25.92	685.26	40.77	688.44	35.97
High	701	5.57	690.17	45.86	693.78	36.77
Total	700.49	27.7	686.98	36.63	692.92	33.55

APPENDIX F

SAMPLE MEANS TABLES, MODEL 3

Table F1

Mean Achievement (Scale Score) for Third Grade Cohort 1999-00 (Fifth Graders), by Mi Use and Gender

MI Use Category	Male		Female		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Not Used	661.92	41.55	676.94	37.29	670.45	39.73
Used	674.4	39.3	669	24.97	671.7	31.17
Total	663.04	41.17	676.39	36.5	670.55	39.02

APPENDIX G

ANOVA RESULTS TABLES, MODEL 4

Table G1

ANOVA for Second Grade Cohort, Dependent Variable: Math Gain from 98 to 99 (=SS99-SS98), Independent Variables: MCS Time, Gender, and Ability Level

Source	*df*	Type IV *SS*	*MS*	*F*	*Eta* Squared
Gender	1	140.622	140.622	0.416	0.004
Ability	2	35066.237	17533.118	51.884*	0.517
Time99	2	1640.677	820.339	2.428	0.048
Gender*Ability	2	1030.064	515.032	1.524	0.03
Gender*Time99	2	817.627	408.814	1.21	0.024
Ability*Time99	3	2482.836	827.612	2.449	0.07
Gender*Ability*Time99	3	1687.366	562.455	1.664	0.049
Error	97	32779.254	337.93		
Total	113	272859			
Corrected Total	112	76622.885			

Note: $N = 113$

*$p < .05$

Table G2

ANOVA for Third Grade Cohort, Dependent Variable: Math Gain from 98 to 99 (=SS99-SS98), Independent Variables: MCS Time, Gender, and Ability Level

Source	*df*	Type IV *SS*	*MS*	*F*	*Eta* Squared
Gender	1	176.724	176.724	0.384	0.003
Ability	2	1.889	0.945	0.002	0
Time99	2	457.949	228.974	0.497	0.009
Gender*Ability	2	1039.729	519.865	1.129	0.02
Gender*Time99	2	227.28	113.64	0.247	0.004
Ability*Time99	3	1086.368	362.123	0.786	0.021
Gender*Ability*Time99	3	1715.375	571.792	1.241	0.032
Error	112	51583.886	460.57		
Total	128	139353			
Corrected Total	127	57289.867			

Note: $N = 128$

Table G3

ANOVA for Fourth Grade Cohort, Dependent Variable: Math Gain from 98 to 99 (=SS99-SS98), Independent Variables: MCS Time, Gender, and Ability Level

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	645.401	645.401	1.429	0.016
Ability	2	5147.934	2573.967	5.700*	0.114
Time99	2	1704.754	852.377	1.888	0.041
Gender*Ability	2	850.031	425.015	0.941	0.021
Gender*Time99	2	167.828	83.914	0.186	0.004
Ability*Time99	3	1691.214	422.803	0.936	0.04
Gender*Ability*Time99	1	826.737	206.684	0.458	0.02
Error	93	40191.474	451.59		
Total	107	73194			
Corrected Total	106	50216.206			

Note: $N = 107$

*$p < .05$

Table G4

ANOVA for Second Grade Cohort, Dependent Variable: Math Gain from 99 to 00 (=SS00-SS99), Independent Variables: MCS Time, Gender, and Ability Level

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	113.121	113.121	0.3	0.003
Ability	2	4739.33	2369.665	6.294*	0.117
Time00	2	948.312	474.156	1.259	0.026
Gender*Ability	2	480.566	240.283	0.638	0.013
Gender*Time00	2	1199.234	599.617	1.593	0.032
Ability*Time00	4	3059.497	764.874	2.032	0.079
Gender*Ability*Time00	4	1347.779	336.945	0.895	0.036
Error	95	35766.29	376.487		
Total	113	131147			
Corrected Total	112	50060.903			

Note: $N = 113$

*$p < .05$

Table G5

ANOVA for Third Grade Cohort, Dependent Variable: Math Gain from 99 to 00 (=SS00-SS99), Independent Variables: MCS Time, Gender, and Ability Level

Source	*df*	Type IV *SS*	*MS*	*F*	*Eta* Squared
Gender	1	300.843	300.843	0.468	0.004
Ability	2	109.71	54.855	0.085	0.002
Time00	2	556.04	278.02	0.433	0.008
Gender*Ability	2	663.219	331.61	0.516	0.009
Gender*Time00	2	482.352	241.176	0.375	0.007
Ability*Time00	3	922.916	307.639	0.479	0.013
Gender*Ability*Time00	2	113.409	56.705	0.088	0.002
Error	113	72625.948	642.708		
Total	128	163445			
Corrected Total	127	78108.867			

Note: N = 128

Table G6

ANOVA for Fourth Grade Cohort, Dependent Variable: Math Gain from 99 to 00 (=SS00-SS99), Independent Variables: MCS Time, Gender, and Ability Level

Source	*df*	Type IV *SS*	*MS*	*F*	*Eta* Squared
Gender	1	590.438	590.438	1.087	0.012
Ability	2	1022.3	511.15	0.941	0.02
Time00	2	858.894	429.447	0.791	0.017
Gender*Ability	2	2295.19	1147.595	2.113	0.043
Gender*Time00	2	2660.009	1330.004	2.448	0.05
Ability*Time00	3	538.791	179.597	0.331	0.011
Gender*Ability*Time00	1	264.974	264.974	0.488	0.005
Error	93	50517.516	543.199		
Total	107	154004			
Corrected Total	106	58183.402			

Note: *N* = 107

APPENDIX H

ANOVA RESULTS TABLES, MODEL 5

Table H1

ANOVA for Second Grade Cohort, Dependent Variable: Math Gain from 99 to 00 (=SS00-SS99), Independent Variables MI Use, Gender, and Ability Level

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	20.52	20.52	0.054	0.001
Ability	2	6429.257	3214.628	8.506*	0.144
NMI00	1	898.707	898.707	2.378	0.023
Gender*Ability	2	227.115	113.557	0.3	0.006
Gender*NMI00	1	390.529	390.529	1.033	0.01
Ability*NMI00	2	2258.101	1129.05	2.988	0.056
Gender*Ability*NMI00	2	343.557	171.779	0.455	0.009
Error	101	38168.826	377.909		
Total	113	131147			
Corrected Total	112	50060.903			

Note: $N = 113$

*$p < .05$

Table H2

ANOVA for Fourth Grade Cohort, Dependent Variable: Math Gain from 99 to 00 (=SS00-SS99), Independent Variables: MI Use, Gender, and Ability

Source	*df*	Type III *SS*	*MS*	*F*	*Eta* Squared
Gender	1	116.727	116.727	0.212	0.002
Ability	2	883.474	441.737	0.801	0.017
NMI00	1	143.742	143.742	0.261	0.003
Gender*Ability	2	765.346	382.673	0.694	0.014
Gender*NMI00	1	924.034	924.034	1.675	0.017
Ability*NMI00	2	1074.846	537.423	0.974	0.02
Gender*Ability*NMI00	2	266.598	133.299	0.242	0.005
Error	95	52414.411	551.731		
Total	107	154004			
Corrected Total	106	58183.402			

Note: $N = 107$

APPENDIX I

SAMPLE MEANS TABLES, MODEL 4

Table I1

Mean Gain (1997-98 - 1998-99) for Third Grade Cohort by MCS Time Category and Ability Group

MCS Time Category	Low Ability		Average Ability		High Ability		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Low	27.63	19.75	26.31	21	31.2	17.65	28.02	19.55
Average	25.18	20.78	24.18	19.28	17.24	28.83	22.81	22.02
High	34.33	22.5	38.5	33.23	**	**	36	23.12
Total	26.86	20.07	25.41	20.1	23.78	24.9	25.32	21.24

** No Subjects

Table I2

Mean Gain (1998-99 - 1999-00) for Third Grade Cohort by MCS Time Category and Ability Group

MCS Time Category	Low Ability		Average Ability		High Ability		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Low	27.92	25.23	28.82	23.62	26.82	27.22	28.1	31.57
Average	22.75	22.29	20.5	20.57	27.53	36.88	22.67	24.55
High	41.25	11.09	15	NA	**	**	36	25.49
Total	27.61	22.66	24.46	22.21	27.16	31.57	25.82	15.17

** No Subjects

Table I3

Mean Gain (1998-99 - 1999-00) for Fourth Grade Cohort by MCS Time Category and Ability Level

MCS Time Category	Low Ability		Average Ability		High Ability		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Low	28.81	19.98	28.91	24.24	30.15	20.27	29.28	22.06
Average	35.58	29.58	17.75	28.06	25	NA	30.76	28.48
High	42.67	40.51	33	NA	**	**	40.25	33.43
Total	32.77	25.52	28.06	24.19	29.96	19.92	29.93	23.43

** No Subjects

APPENDIX J

SAMPLE MEANS TABLES, MODEL 5

Table J1

Mean Gain (1998-99 - 1999-00) for Second Grade Cohort by MI Use Category and Ability Level

MI Use Category	Low Ability		Average Ability		High Ability		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Not Used	16.2	23	18.86	16.67	33.62	20.88	23.79	21.56
Used	12.45	12.33	36	18.55	40	18.12	31.15	19.95
Total	14.87	19.72	28.34	19.54	35.26	20.15	26.79	21.14

Table J2

Mean Gain (1998-99 - 1999-00) for Fourth Grade Cohort by MI Use Category and Ability Level

MI Use Category	Low Ability		Average Ability		High Ability		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Not Used	27.95	22.03	29.09	24.25	31.13	15.7	28.97	22.35
Used	42.9	30.4	25.31	24.82	29.5	21.73	31.35	25.15
Total	32.77	25.52	28.06	24.19	29.96	19.92	29.93	23.43

VITA

VICKI C. KIRK

Education:	February 2000 – successfully completed qualifying examination and formally accepted as a doctoral candidate at East Tennessee State University in the Educational Leadership and Policy Analysis Program; presently completing the dissertation process
	1991 graduate of East Tennessee State University with Master of Arts in Teaching; gpa – 4.0
	1983 graduate of East Tennessee State University with B.S. in Medical Technology, summa cum laude (valedictorian)
	Professional Experience: June 2000 to present: Assistant Superintendent of Instruction – Greeneville City Schools, Greeneville, TN
	August 1999 – June 2000: Assistant Principal – Greeneville Middle School, Greeneville City Schools, Greeneville, TN
	August 1992 to July 1999: Eighth Grade Science Teacher - Greeneville Middle School, Greeneville City Schools, Greeneville, TN
	September 1983 to August 1992: MEDICAL TECHNOLOGIST – Laughlin Memorial Hospital, Greeneville, TN
Activities and Honors:	Visited Singapore schools as a member of a cadre of teachers from Tennessee and presented to school and community groups
	Served as a member of the Character Development Team; presented to the Tennessee School Board Association, to the Governor's Conference for Safe and Drug Free Schools, and to Senator Bill Frist
	Awarded Greene County Partnership's Educator of the Year, 1996