THE EFFECTS OF PERSONALIZED PRACTICE SOFTWARE ON LEARNING MATH STANDARDS IN THE THIRD THROUGH FIFTH GRADES

by

ANGELA NICOLE GOMEZ

A DISSERTATION

Presented to the Department of Educational Methodology, Policy, and Leadership
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Education

December 2012

DISSERTATION APPROVAL PAGE

Student: Angela Nicole Gomez

Title: The Effects of Personalized Practice Software on Learning Math Standards in the Third through Fifth Grades

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Education degree in the Department of Educational Methodology, Policy, and Leadership by:

Dr. Gerald Tindal	Chair
Dr. Julie Alonzo	Member
Dr. Kathleen Scalise	Member
Dr. Arthur Farley	Outside Member

and

Kimberly Andrews Espy	Vice President for Research and Innovation/Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded December 2012

DISSERTATION ABSTRACT

Angela Nicole Gomez

Doctor of Education

Department of Educational Methodology, Policy, and Leadership

December 2012

Title: The Effects of Personalized Practice Software on Learning Math Standards in the Third through Fifth Grades

The purpose of this study was to investigate the effectiveness of *MathFacts in a Flash* software in helping students learn math standards. In each of their classes, the third-, fourth-, and fifth-grade students in a small private Roman Catholic school from the Pacific Northwest were randomly assigned either to a control group that used flash cards and worksheets or to a treatment group that used a computer software program to practice grade-level appropriate math facts. Students advanced to math facts at the next grade level after completing the levels appropriate to their own. A crossed design allowed the two groups of students in each of the grades to participate in their respective intervention and control treatments over the course of 6 weeks before they received the alternative treatment. Students took equivalent forms of curriculum-based measures for their grade level at the beginning, middle, and end of the study (e.g., third graders took third grade assessments) and equivalent forms of curriculum-based measures at the middle and end of the study for the next grade level (e.g., third graders took fourth grade assessments).

A correlated-groups *t*-test was conducted to determine the significance of the computer software program on students' performance on the grade-level measures, and an independent-groups *t*-test was conducted to determine the significance of the computer software program on students' performance on the subsequent grade-level measures. The results of the study indicate that there was not a significant difference in math scores between students practicing math facts with *MathFacts in a Flash* and those practicing math facts with flash cards and worksheets in both the on-grade and subsequent-grade-level measures. The findings are discussed in the context of the ways computer software may still be used to increase student proficiency with learning math standards in the third, fourth, and fifth grades.

CURRICULUM VITAE

NAME OF AUTHOR: Angela Nicole Gomez

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR
University of Portland, Portland, OR
University of Saint Mary, Leavenworth, KS

DEGREES AWARDED:

Doctor of Education, 2012, University of Oregon
Master of Education, 2004, University of Portland
Bachelor of Science, Elementary Education, 2002, University of Saint Mary

AREAS OF SPECIAL INTEREST:

Instructional Technology
Special Education
Music

PROFESSIONAL EXPERIENCE:

Kindergarten Teacher, Saint John the Baptist, Milwaukie, OR, 2007-present

Teacher, Our Lady of Sorrows, Portland, OR, 2002-2007
First and second grade combination and music (K-4), 2006-2007
Third grade 2002-2005

PUBLICATIONS:

Gomez, A. (2011). *Organizing teacher collaboration between schools for student achievement*. Retrieved from http://hdl.handle.net/1794/11119

ACKNOWLEDGMENTS

I want to thank all of my friends and family who supported me during the process of completing this dissertation. The road was a little smoother and a lot shorter knowing that I had such caring, compassionate, and patient people helping me along the way! I want to thank the professors at the University of Oregon who helped me complete the program, including my advisor and committee members. I am fortunate to have the opportunity to complete an advanced degree program, and I know that it would not have been possible to do on my own. I am extremely thankful to God for the gifts I have received through the goodness of others.

Dedicated to the Gomez and Avila matriarchs who are always with me in spirit.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

Attention to the importance of standards in mathematics education has increased since the release of the Common Core State Standards (CCSS) in 2010 (Achieve, 2010). The CCSS in math are similar to the standards assessed on the National Assessment of Educational Progress (NAEP) mathematics assessment. Five content areas are emphasized: (a) number properties and operations; (b) measurement; (c) geometry; (d) data analysis, statistics, and probability; and (e) algebra. An important component of mathematics that is required in each of these five content areas is computation (NAEP, 2010). Yet, only a few studies have examined the relationship between math computation and other math content areas, showing that there is a relation between arithmetic and algorithmic computation as well as arithmetic and arithmetic word problems (Fuchs et al. 2006), and that students' difficulty in computation does not necessarily predict difficulty in problem solving and vice versa (Fuchs et al. 2008).

Even though the research on the relationships between math content areas is limited, there is no doubt that computation skills play a role in students' success in mathematics. In fact, the National Mathematics Advisory Panel ([NMAP], 2008) stressed the importance of computational fluency with whole numbers and the necessity of sufficient and appropriate practice to develop automaticity. Moreover, mathematics researchers determined that students need to practice basic facts for about ten and no less than five minutes a day for fluency, especially for students who are struggling in mathematics (Gersten et al., 2009).

Although the CCSS provide educators with critical information about the content that students need to learn to be successful, they do not provide information or guidelines about how this content should be taught. Educational leaders and teachers do this in many different ways. In mathematics, instructional recommendations come from various professional groups and agencies. The National Council of Teachers of Mathematics ([NCTM], 2009) suggests that incorporating technology into mathematics teaching is an important and viable instructional practice for teachers to consider:

> Technology is essential in teaching and learning mathematics; it influences the mathematics that is taught and enhances students' learning. Students can develop deeper understanding of mathematics with the appropriate use of technology….The existence, versatility, and power of technology make it possible and necessary to reexamine what mathematics students should learn as well as how they can best learn it. (p.3)

The Institute of Education Sciences further recommended the use of technology-based supplemental programs (Gersten et al., 2009; NMAP, 2008) because contemporary mathematics curricula do not emphasize math fact practice for fluency (Gersten et al., 2009). When implemented with fidelity, high-quality Computer Assisted Instruction (CAI) can be considered a useful tool in developing fluency (NMAP, 2008). Even though there are recommendations for the use of technology in teaching and learning math, a surprising 61% of fourth-grade students report that they *never or hardly ever* use computers for math at school (NAEP, 2009).

However, the National Mathematics Advisory Panel (2008) also reported that the nature and strength of the effectiveness of instructional software vary from study to study, and that there is insufficient research available to help educators identify the factors that impact the effectiveness of instructional software in mathematics.

To explore whether there is a positive relation between students' use of computers during mathematics instruction and their learning of math standards, I searched for relevant studies from the electronic databases referenced in Table 1. After the initial search for peer-reviewed journals on ERIC, I further limited the results to elementary school age populations and focused on meta-analyses. I obtained copies of studies from the most recent meta-analysis (Li & Ma, 2010) that included the use of tutorials in third through fifth grades.

Table 1

Literature Search Process

Search Engines and Sites	Keywords	Number of Articles Found
ERIC	Mathematics *and* Computation *and* Technology	200
PsycINFO	Mathematics *and* Technology *and* Meta-analysis	24
ISI Web of Knowledge	Mathematics *and* Technology *and* Meta-analysis	10
SAGE Reference Online	Mathematics *and* technology	10

I present the literature in a funnel technique (starting broadly and ending narrowly) to show the complexity and variety of computer research in education that has impacted the progression of its use over time. I begin with an examination of whether students' computer use is associated with positive learning outcomes in general, and then specifically with positive learning outcomes in mathematics. I then define and examine the use of tutorial software in learning basic mathematics skills related to best practices in differentiation or personalization. Next, I explore how previous researchers have used different measures to determine whether tutorial software has had a positive impact on student learning outcomes. Finally, I propose a study that incorporates student use of personalized computer software to positively affect how students learn math standards in the third, fourth, and fifth grades.

Computers in Education

The evolution of computers in education is broad and narrow at the same time; broadly encompassing many different content areas, and at its start, narrowly focused on drill and practice. Early technology implementation evolved from behavioral based theories (Skinner, 1989), often called Computer Assisted Instruction (CAI) or Computer Based Instruction (CBI), whereby students practiced specific educational content, received feedback on their performance, and either repeated or moved to the next step in the learning progression (Means, 2008). Over time, technology became more open-ended and complex, focusing on student understanding and internal cognitive processing (Means, 2008). Meta-analyses and reviews of evaluation reports over time on the many content areas researched provide evidence that it is necessary to evaluate the effects of technology on student learning experiences. Even though the following analyses were published close in time, they were quite different in emphasis: One explored studies over a wide time span (Kulik & Kulik, 1991), and the other specifically examined the use of computer-based systems popular in schools at the time (Becker, 1992).

The Kulik and Kulik (1991) meta-analysis of 254 studies examined the effects of computer-based instruction (CBI) on student performance in different content areas from 1966-1986. They found that CBI had a moderate but significant effect size (ES = .30) on student performance. Effect sizes were greater in those studies that were published in journals (ES = .44), that had different teachers for the control and treatment groups (ES = .39), and that had intervention durations of 4 weeks or less (ES = .42). Additionally, they found that CBI reduced the time needed for instruction and had small but positive effects on students' attitudes toward instruction, coursework, and computers.

In response to districts' desire for empirical evidence on the effects of computer-based learning systems in schools, vendors began to produce evaluation reports on their products (Becker, 1992). Becker (1992) analyzed a collection of 30 evaluation reports to determine the effect size of computer-based integrated learning systems in the elementary and middle school grades. He defined integrated learning systems as, "networked comprehensive basic skills software from a single vendor" (p. 1). Although Becker found that integrated learning systems had a moderate but positive effect on student achievement, he reported that there were weaknesses in many of the designs and evaluations. He cautioned that districts purchasing integrated learning systems need to read the evaluation reports and consider who evaluated the program, who the participants were, what achievement test was used to determine the effects of the program, how much information was given about implementation for the control and test groups, and how the data were analyzed.

Overall, Kulik and Kulik (1991) and Becker (1992) reported that incorporating computer-based learning into instruction had moderate but positive effects on student achievement in various subject areas. In addition, they also demonstrated the importance of considering the variables, participants, and study designs in each study. Research on the effects of mathematics computer-based learning on student achievement mirrors these findings.

Computers in Mathematics Education

Just as research on computers in general education is broad and narrow, so too is research on computers in mathematics education. In an effort to provide effect sizes for the impact of computer-based instruction on student achievement in mathematics, researchers categorize a broad range of studies together based on certain features. This categorization narrows the information, but can make it difficult to determine other features of the studies which may or may not have had an impact on student achievement (NMAP, 2008). Nevertheless, it is still worthwhile and necessary to consider the overall effects from a meta-analysis as opposed to placing a significant amount of importance on the findings of a single study (Kulik & Kulik, 1991).

Kulik (2003) has reviewed studies involving integrated learning systems, specifically in mathematics, since 1990. He found that effect sizes were large enough to be considered educationally meaningful, but also suggested ways to strengthen future research in this area. Suggestions included devoting more time per week to the integrated learning system instruction, combining regular classroom instruction with integrated learning instruction, and allowing students to work in pairs rather than individually.

In examining how teachers deliver mathematics content, Slavin and Lake (2008) reviewed 87 studies in which they determined the effectiveness of mathematics curricula (related to textbooks), computer-assisted instruction, and instructional process programs in elementary school mathematics. Though the effects of computer-assisted instruction were moderate, they were more effective than mathematics curricula. The authors determined that instructional processes such as cooperative learning had the strongest positive effects, but suggested that a combination of the three areas would create a stronger instructional math program.

Finally, the most recent meta-analysis was published by Li and Ma (2010), who also focused on the effectiveness of computer technology for student learning in mathematics. They defined computer technology as software rather than hardware in four main areas: (a) tutorial, such as games and drill for practice; (b) communication media, such as email and video-conferencing; (c) exploratory environments, such as simulations and hyper-media based learning; and (d) tools, such as spreadsheets and instructional management software. Students' scores on solving mathematical problems on mathematics tests, standardized or teacher/researcher-made, were used as indicators of student mathematics achievement. Using these definitions, Li and Ma (2010) narrowed their analysis down to 46 primary studies conducted after 1990 in kindergarten through grade 12 settings. Overall, they found that there was a moderate, but significantly positive effect of computer technology on mathematics achievement. See Table 2 for a comparison of the study characteristics that had larger effects than others.

Interestingly, the effects of computer technology were similar for various study characteristics or components. The four computer technology formats shared the same level of effectiveness, meaning that one did not prove to be more effective than the others. Li and Ma (2010) also found that students of both genders and all races and socio-economic status levels benefitted equally from the implementation of computer technology in the classroom. There was also little to no difference between the studies that used rigorous scientific methods (random assignment) and those that were less rigorous.

Table 2

Li and Ma (2010) Study Features and Effect Sizes

Study Feature	Characteristics Compared			
	Group	*ES*	Group	*ES*
Special education status	Special education students	1.33	General education students	0.80
Method of teaching	Constructivist approach	1.20	Traditional approach	0.83
Publication year	Before 1999	1.04	After 1999	0.64
Level of education	Elementary	0.93	Secondary	0.87
Study duration	One term	1.00	Longer than 1 year	0.64
Type of assessment	Teacher/researcher-made	1.02	Standardized	0.78

The findings of Kulik (2003), Slavin and Lake (2008), and Li and Ma (2010), are comparable to the findings of Kulik and Kulik (1991) and Becker (1992): Students make gains in their learning that are small but significant when teachers integrate computers into the curriculum. Considering the current focus on the effectiveness of computer software on learning basic math facts, what components of the software specifically impacted student achievement?

Defining and Examining Tutorials: The Importance of Differentiation

Given that the National Mathematics Advisory Panel (2008) has cautioned the use of computer software due to the inconsistencies across available studies, it is important to unpack the variables in the studies to find commonalities. I examined how tutorials were used in the studies included in the meta-analysis conducted by Li and Ma (2010). I focused specifically on those studies that included students in grades 3 through 5.

Li and Ma (2010), following the recommendation of Means (1994) and Lou, Abrami, and d'Apollonia (2001), grouped studies in their meta-analysis into four major categories. *Tutorials* included programs that explicitly taught mathematics to students through a stimulating environment that provided information, demonstration, drill, and practice (Lou et al., 2001). The tutorial category also included computer-assisted instruction (CAI), math games, and drill and practice software. In their meta-analysis, Li and Ma (2010) defined CAI as direct instruction programs or drill and practice. Most of the tutorials provided some form of differentiation for students and were interactive.

Tomlinson (1999) defines differentiation as a teacher's response to a student's needs. The teacher can respond to a student's need for something different in content, process, product, or learning environment based on the student's readiness, interest, and profile. The content refers to what students will learn, the process to the activities students will complete, the product to how students will demonstrate their learning, and the learning environment to the conditions that create an atmosphere for learning. Educators and instructional programs can employ multiple methods and strategies to differentiate the content, process, product, and/or learning environment within which instruction occurs.

Differentiating process. Several of the software programs I examined, for example, allowed students to set individual goals and work toward mastery of their goals (Blanton, Moorman, Hayes, & Warner, 1997; Irish, 2002; Martindale, Pearson, Curda, & Pilcher, 2005; Wittman, Marcinkiewicz, & Hamodey-Douglas, 1998; Xin, 1999). The individualization of goal-setting included playing against the computer for accuracy with unlimited time (Xin, 1999), and self-goal setting for accuracy and fluency (Blanton et al., 1997; Irish, 2002; Martindale et al., 2005; Wittman et al., 1998). This is synonymous with differentiation of process (Tomlinson, 1999). All participating students worked toward a goal, but the rate and accuracy of their goal was based on their own sense of readiness.

Differentiating content. As a prerequisite to goal-setting, a few of the studies (Irish, 2002; Wittman et al., 1998) incorporated the use of compacting (Tomlinson, 1999) whereby each student began the program with a pretest that determined what the student knew. The program recorded any difficulties that arose for the student, and then provided targeted practice or a more challenging sequence of practice based on their initial performance. In one study (Wittman et al., 1998), the researchers guided student content rather than the computer program.

Differentiating the learning environment. Another important component of differentiation includes students' engagement with computer software that allows them to interact with the program as well as be exposed to animated graphics (Martindale et al., 2005; Xin, 1999) and sounds (Blanton et al., 1997; Irish, 2002; Xin, 1999). Some computer software programs also provided students with the opportunity to receive feedback for correct and incorrect responses (Irish, 2002), with some providing detailed explanations for the correct or incorrect answers (Martindale et al., 2005; Xin, 1999). These interactive features are similar to the learning environment aspect of differentiation (Tomlinson, 1999). The visual and audible feedback from the programs creates the conditions that set the atmosphere for learning (Tomlinson, 1999).

Differentiating product. Because the nature of tutorial software is to provide students with information, demonstration, drill, and practice (Lou et al., 2001), there was little evidence of differentiation of product (the way in which students demonstrated their learning) (Tomlinson, 1999). However, one study (Blanton et al., 1997) incorporated a program that required students to produce different multi-media and physical representations of their learning as they worked through the program.

Even though it is important to determine the specific characteristics of the computer software that seem to make a difference in its effectiveness, it is also important to reflect on the measurements given to students in the studies. Becker (1992) addressed the need to analyze the appropriateness of the measures to determine the effects of the program. What types of measurements were used in the studies included in the Li and Ma (2010) meta-analysis focused on students using tutorials in the third through fifth grades?

Comparing and Contrasting Tests and Designs

Not only does assessment play a large role in determining the effects of any treatment, it plays a role in teaching and learning in classrooms:

> When assessment is an integral part of mathematics instruction, it contributes significantly to students' mathematics learning. Assessment should inform and guide teachers as they make instructional decisions. The tasks teachers select for

> assessment convey a message to students about what kinds of mathematical knowledge and performance are valued. Feedback from assessment tasks helps students in setting goals, assuming responsibility for their own learning, and becoming more independent learners. (NCTM, 2009, p. 2)

Li and Ma (2010) determined that the use of non-standardized tests produced larger effects than standardized tests. I identified the types of tests used in the studies involving students in the third through fifth grades who received mathematics instruction via tutorials. I found that the tests were mostly criterion-referenced, and that most studies used a pre-test post-test design to determine whether the programs implemented had positive effects on students' mathematics achievement.

Norm-referenced vs. criterion-referenced tests. Two of the eight studies I analyzed focused on norm-referenced tests (Wittman et al., 1998; Xin, 1999). Norm-referenced tests are designed to compare a student with other students of the same age or grade level. They are standardized, so they have specific criteria for administration and scoring that can help eliminate subjectivity in interpreting results given the criteria were followed (Overton, 2000). The remaining six studies used criterion-referenced tests. Criterion-referenced tests may or may not be standardized and are designed to determine an individual student's mastery of specific criteria (Overton, 2000). Of the criterion-referenced group, several studies employed the use of assessments that closely matched district, state, or national tests (Blanton et al., 1997; Martindale et al., 2005; Quinn & Quinn, 2001; Royer, Greene, & Anzalone, 1994; Salerno, 1995). In another study, the author created a test that focused on specific math facts that students practiced and that was designed to measure students' use of mnemonic devices they had been taught throughout the study (Irish, 2002).

Pre-test post-test vs. multiple baseline designs. The use of multiple baseline designs was limited to one study (Irish, 2002). Multiple baseline or time-series designs use more than one set of data collected over time to make comparisons between groups (Babbie, 2010). Irish (2002) tested students weekly (over 18 weeks) to determine the impact on math fact recall of learning mnemonic devices through computer software. The remaining seven authors (Blanton et al., 1997; Martindale et al., 2005; Quinn & Quinn, 2001; Royer et al., 1994; Salerno, 1995; Wittman et al., 1998; Xin, 1999) designed their studies to include a pre-test and a post-test. In a pre-test post-test design, students are tested prior to and then after receiving the treatment. The differences between the first and last test are used to help explain the impact of the treatment (Babbie, 2010).

The implications of the tests and how they were administered, to a certain point, reflect common uses of assessment today and shed light on areas of consideration for the future. The primary use of criterion-referenced tests over norm-referenced tests coincides not only with the NCTM's (2009) recommendation to use assessment to help teachers make instructional decisions, but also with the National Mathematics Advisory Panel's (2008) recommendation to focus teaching, learning, and assessment on specific items to prepare students for success in algebra. The criterion-referenced tests are designed to highlight a student's achievement on specific content rather than their achievement compared to other students (Overton, 200). How can we take the information regarding differentiation, testing, and design and use it to direct the future of research on computer software, basic skills, and assessment?

Learning From the Research and Moving Forward

It is critical to be cognizant of the limitations that occur in studies and to consider all of the variables that may or may not impact outcomes and future research (Becker, 1992; National Mathematics Advisory Panel, 2008). The tutorial studies analyzed by Li and Ma (2010) contained many characteristics found in Table 3 that do not apply to the question of my research study focusing only on math facts. Specifically, do third-, fourth-, and fifth-grade students who use *MathFacts in a Flash* software to practice basic math facts score significantly higher on mathematics curriculum based measures than third-,fourth-, and fifth-grade students who use flashcards and worksheets for practice?

Table 3

Study Characteristics That Do Not Align With the Research Question

Characteristic	Quality and Author
Setting	After school program (Blanton et al., 1997); Summer school program (Quinn & Quinn, 2001)
Focus of Study	Mathematics anxiety only (Wittman et al., 1998); Mnemonic devices (Irish, 2002); Multiple areas: decimals, fractions, problem solving (Xin, 1999) addition, subtraction, percentages (Salerno, 1995); Reading and math (Martindale et al., 2005; Quinn & Quinn, 2001; Royer et al., 1994); Unable to determine (Blanton et al., 1997)

Even though there were several characteristics that were not applicable to the current study question, the literature review revealed areas for consideration, replication, and alteration. Because there is research on the strength of a short intervention time (Kulik & Kulik, 1991; Li & Ma, 2010), my study will take place over the course of a term (12 weeks). In addition, over the course of a week, students will spend about ten minutes and no less than five of practice of basic facts for fluency (Gersten et al., 2009), and when using an integrated learning system, students will spend the recommended time on the program, not less (Kulik, 2003). Students will have learning goals to work toward mastery (Blanton et al., 1997; Irish, 2002; Martindale et al., 2005; Wittman et al., 1998; Xin, 1999), and they will be engaged in computer software that is interactive to provide them feedback on their progress toward the learning goals (Irish, 2002; Martindale et al., 2005; Xin, 1999).

Students will take criterion-based measures or curriculum-based measures. Reliable measures of computational knowledge include curriculum-based measures (Thurber, Shinn, & Smolkowski, 2002). These measures were used to determine the effects of basic skills practice with technology for transfer to multiple areas of mathematics (Blanton et al., 1997; Martindale et al., 2005; Quinn & Quinn, 2001; Royer et al., 1994; Salerno, 1995).

Based on these considerations and to determine how these findings might apply to a single school with third-, fourth-, and fifth-grade students, I propose a study in which I hypothesize the following effects on mathematics curriculum-based measures:

- H_0: There is no significant difference in math scores between students practicing math facts with *MathFacts in a Flash* and those practicing math facts with flash cards and worksheets.
- H_a: There is a small but significant increase in math scores for students practicing math facts with *MathFacts in a Flash* compared to students who practice math facts with flash cards and worksheets.

CHAPTER II

METHODOLOGY

In this section, I present the populations from which I sampled students. I describe a convenience sample in a small Catholic school. Next, I describe the measures used to document outcomes of frequent practice with math facts. Then, I explain the treatment conditions that were used to provide students with fluency practice with basic math facts: a computer-based software package and a traditional paper-pencil approach. I then describe the operational procedures in which teachers and students were oriented to participate in a 12-week study. I conclude with a description of how I analyzed the data.

Sampling

The Archdiocese of Portland, at 29,717 square miles, spans the western part of the state of Oregon from the summit of the Cascades to the Pacific Ocean. The Archdiocese oversees 41 elementary schools with approximately 8, 860 students, 74% of whom are identified as White and 83% of whom are identified as Catholic. Students of Hispanic and Asian backgrounds are the largest minority groups at 7.7 % and 7.5% respectively. The average elementary school enrollment is 216 students (Bunce, 2010).

The percentage of minorities in Catholic schools across the country has increased from 10.8% in 1970 to 29.8% in 2010 (McDonald & Schultz, 2010). During this 40-year period, Hispanic enrollment has increased by 12.8%, Black/African American enrollment by 7.5%, Asian American enrollment by 4.5%, and the enrollment of multiracial students by 3.7% (McDonald & Schultz, 2010). Non-Catholic student enrollment has also increased from 2.7% of the population in 1970 to 14.5% of the population in 2010 (McDonald & Schultz, 2010). The percentage of lay (not part of the clergy) men and women teaching in Catholic schools has increased in the last decade from 85% to 96.3% (McDonald & Schultz, 2010).

Student demographics. The students in the study attended *Saint John the Baptist School* in Milwaukie, OR, a Catholic school in the Archdiocese of Portland. Saint John the Baptist is a private school consisting of students coming from mostly Caucasian middle and upper class families. The school had one class for each grade, Pre-Kindergarten through eight, with a total enrollment of 206 students. Four students were identified as having a disability. These students received services for an Instructional Service Plan (ISP) at different public schools during the school day. Enrollment and religious and ethnic composition of the school are presented in Tables 4 and 5.

Table 4

Saint John the Baptist 2011-2012 Enrollment by Gender and Grade

Grade	Male	Female	Total
Pre-K	7	7	14
K	10	10	20
1	7	9	16
2	11	11	22
3	8	14	22
4	10	14	24
5	7	11	18
6	13	7	20
7	13	13	26
8	6	18	24
Total	92	114	206

Table 5

Saint John the Baptist 2011-2012 Enrollment by Religion and Ethnicity

Ethnicity	Catholic	Non-Catholic	Total
Black	0	1	1
Hispanic	3	0	3
Asian	2	5	7
Hawaiian/Pacific Islander	4	0	4
Native American/ Native Alaskan	0	0	0
White	151	39	190
Unknown		1	1
Total	160	46	206

Teacher demographics. The school employed one teacher per grade, a physical education teacher, a music teacher, a librarian, and a reading specialist. Of the faculty, 13 were female and 2 were male. The religious and ethnic composition of the teaching faculty was similar to that of the students: The majority of teachers identified themselves as White and Catholic. Two teachers were religious sisters and the remaining were lay men and women. The faculty averaged eight years of teaching experience and four years of employment at the school.

Participant demographics and experience. In the winter of the 2011-2012 school year, the third, fourth, and fifth grade classes participated in the study. No students in these classes were identified as having any disabilities. I obtained written consent from the students and their parents prior to the study to include the students' scores in the data analysis. Two students in the third grade and three students in the fourth grade did not give consent for the use of their scores in the study. One student from fourth grade changed schools during the first week of the study. The demographics for the students from each of the three classes who participated in the study are provided in Table 6.

Table 6

Saint John the Baptist 2011-2012 Descriptive Statistics of Study Participants

Grade	Female	White	Hispanic	Asian	Other
Third (*n*= 20)	13	15	2	2	1
Fourth (*n*= 20)	11	18	0	1	1
Fifth (*n*= 18)	11	15	0	2	1

The third grade teacher was in her third year of teaching, the fourth grade teacher was in her second year of teaching, and the fifth grade teacher was in her fifth year of teaching. The fourth grade teacher went on maternity leave at the start of the fifth week of the study and her long-term substitute was a first year teacher. All teachers conducted their math classes in a 60-minute block five days a week in each of their classrooms. The first 10 minutes of the class was dedicated to math fact practice and the remaining 50 minutes was spent on introducing, practicing, and reviewing new mathematical concepts.

Measures

Students completed a series of math curriculum-based measures from the easyCBM and AIMSweb formative assessment systems at the beginning, middle, and end of the study (weeks 1, 6, and 12). The two measures focus on separate areas for analysis. The easyCBM measures focus on math concepts because the research question if focused on learning math standards and the AIMSweb on math computation because research suggests that fluency in math computation is important for success in all areas of mathematics (National Mathematics Advisory Panel, 2008).

easyCBM system. easyCBM is a benchmark and progress monitoring system that includes benchmark tests that can be administered three times a year (e.g., in the fall, winter, and spring), and progress monitoring tests that can be administered between the benchmarking periods. easyCBM math tests were designed to focus on student conceptual understanding of the National Council of Teachers of Mathematics Focal Point Standards (Numbers and Operations, Measurement and Data Analysis, and Numbers and Operations and Algebra) rather than computational skills (Alonzo & Tindal, 2010). A screening test consists of 45 items covering all of the Focal Points and there are 10 alternate forms with 16 items that assess content aligned with the three Focal Points for a total of 30 progress monitoring tests (Alonzo & Tindal, 2010).

The 10 progress monitoring tests in each of the three focal points were designed to be of equivalent difficulty. As described in Technical Reports No. 0901, 0902, and 0903 (Alonzo, Lai, & Tindal, 2009), a Rasch model and distractor analysis were used to determine the difficulty and appropriateness of all items that would be included in the item bank for students in a given grade level. Once all of the items were either retained for or rejected from the item bank, based on the Rasch model and distractor analysis, the alternate forms of the tests were created. Each alternate form was designed so that the items progress in difficulty from easy to more difficult; however, the most difficult item, the sixteenth item, comes in place of item five, leaving the fifth easiest item for the last on the test. This design feature was intended to provide more information for teachers on students' efforts. If students are able to correctly answer items six, seven, and eight, but miss item sixteen, they might have stopped trying by the end. The last item would have been easier than the preceding items (Alonzo & Tindal, 2010). The alternate forms were designed to be given no more than once every three weeks.

Internal consistency is reported as α = between .70 and .80, .80s, and between .80 and .90 for grades 3, 4, and 5, respectively. Split-half reliability is reported as between .50 and .80, .70s, and between .70 and .80 for grades 3, 4, and 5, respectively (Nese et al., 2010).

No teacher training is required for test administration, as students take the math tests individually on the computer. The teacher may also print out the test to administer to the students in a group. Teachers are responsible for ensuring that students know how to log-in by typing in a username, selecting their name from a drop-down list, and choosing the correct test. Teachers must also ensure that the test-taking environment is quiet, that students have access to scratch paper and a pencil or any other accommodations they need, and that students are focused on their own test. If students are taking the test using the computer-based platform, the program brings the student back to the last item on which they were working in the event that they are timed-out.

Students' responses are scored dichotomously, where by each correct response earns a score of 1 and each incorrect response earns a score of 0. A sum of the points earned for the measure is calculated and provided as a total score for the teacher. In the event that the test is not taken online, the teacher enters the students' answer choices into the computer for the computer to score (Alonzo & Tindal, 2010). See Appendix B for screen shots, reporting pages, and examples of the easyCBM program.

AIMSweb system. Students also completed a series of curriculum-based measures in math from AIMSweb (Pearson Education, 2008), a benchmark and progress monitoring system that also includes benchmark tests intended to be administered three times a year (e.g., in the fall, winter, and spring), and progress monitoring tests can be administered as needed between the benchmarking periods.

AIMSweb provides Mathematics Computation (M-COMP) probes for Grades 1-8 with 30 alternate forms per grade. Each probe consists of 2 pages of computational problems printed front to back that are arrayed in rows with boxes around them. Students are given eight minutes to complete as many problems as possible. Third grade problems consist of column addition, basic facts, and complex computation. Fourth grade problems consist of basic facts, complex computation, fractions, and decimals. Fifth grade facts consist of basic facts, complex computation, fractions, decimals, reducing, percentages, and conversions (Pearson Education, 2010).

The probes at each grade level (1-8) were designed to be of equivalent difficulty. An anchor probe was developed for each grade from items that were field-tested and then evaluated based on point-biserial correlations and item difficulty. Each equivalent probe was constructed to replicate the item type proportions, difficulty, and item placement on the anchor probe. Easier items were generally placed at the beginning of each probe to increase the amount of data collected from at-risk learners. Internal consistency is reported as $\alpha = .89$, .87, and .91 for grades 3, 4, and 5, respectively. Split-half reliability is reported as .90, .91, and .93 for grades 3, 4, and 5, respectively (Pearson Education, 2010).

AIMSweb M-COMP is a standardized test that can be administered in individual, small group, or whole class settings. Teachers must read the instructions that accompany the test to the students verbatim, following the instructions for how to direct students when they lose interest or have questions. Each probe includes an answer key. The student is given one point for the correct answer and zero points for an incorrect answer. All points are added to determine the final score (Pearson Education, 2010). See Appendix C for an example test and an example answer key for scoring.

Treatments

Students were exposed to two different conditions during the 12 weeks of the study. The conditions were counterbalanced so that for six weeks, half of the students participated in the treatment condition first and then for the remaining six weeks, they participated in the control condition. The reverse is true for the other half who first participated in the control condition. The treatment condition consisted of math fact practice with a software program called *MathFacts in a Flash*, and the control condition consisted of math fact practice with flash cards and worksheets.

Treatment condition. *MathFacts in a Flash* software provides students in first through sixth grades with practice in developing automaticity of addition, subtraction, multiplication, and division facts. Students completed a 40-item two-minute timed test on the computer for each new math level. There are 62 levels of practice that target specific facts in the areas of addition, subtraction, multiplication, and division, and that provide practice with mixed facts review. The program provides immediate on-screen feedback to students with their time and accuracy information, and showed any missed problems.

Students then practiced known and unknown math facts, including those they missed on the pre-test on MacBook Air laptops. When they were ready (they decided when they thought they could pass another test), they took another 40-item timed test until all problems were answered correctly in the time goal. The time goals were pre-set for 2 minutes but could be changed by the teacher for individual students. The program advanced students to the next level in their sequence when they met their goal. For example, students meeting the 2-minute goal for addition with 0 and 1 would advance to practice with items requiring addition with 2 and 3. Teachers also had the ability to adjust the sequence of levels for individual students, with the ability to track and print out student progress reports. Parents could be notified of student progress through emails, and students could practice at home with a log-in; however, their practice at home did not count toward their school goals (Renaissance, 2009). See Appendix A for *MathFacts in a Flash* screen shots, student report examples, and the scope and sequence used for each of the three grades.

During the 2002-2003 school year, 4,224 elementary and secondary students from 13 schools in 10 states practiced their math facts using *MathFacts in a Flash* software. Students practiced anywhere from 5–15 minutes daily, and teachers were asked to monitor practice and provide appropriate instruction. Researchers used an efficiency indicator to adjust for differences in program usage. They found that all grade levels with higher scores on the efficiency indicator had greater gains in achievement. In response to a survey given at the end of the study, 59% of students said they liked math better after using *MathFacts in a Flash*, and 93% of teachers said that the program helped their students become better at math (Ysseldyke, Thill, Pohl, & Bolt, 2005).

Control condition. The students in the control condition completed a worksheet of 40 math facts focused on a specific fact in two minutes and needed to answer all questions correctly in order to advance to the next level. Students in the control condition followed the same scope and sequence as the students in the treatment condition. If a student did not pass the 40 question test in two minutes, then they individually practiced their math facts with front and back flash cards (the question was on one side and the answer on the back). The flash cards were bagged in sets according to the levels the students needed. Students received no assistance on how to use the flash cards apart from reminders that they needed to work independently. As the teacher monitored students during their practice time, students let the teacher know when they were ready to take a test. Two minutes before the end of the practice session, those who wanted to test retrieved the corresponding math sheet from a file folder in a crate of tests and the teacher timed them. The teacher graded the worksheets after class and gave their papers back at the start of math class the next day.

Design

I implemented a crossed design in my study in which students participated in all conditions. The purpose of the analysis was to determine if the treatment effects were significant, or the result of chance given the small sample size (Stockburger, 1996).

In January 2012, half of the third, fourth, and fifth grade students, in each of their respective classes, at Saint John the Baptist Catholic School were assigned to the Renaissance *MathFacts* treatment condition (RMF) and half were assigned to the paper-pencil control condition (PP). Students were rank-ordered by their pre-test easyCBM scores and alternately placed in either the RMF condition or the PP condition so that there was an equal representation of students in each condition who scored high and low. At the end of the sixth week of the study, students in the RMF treatment condition crossed over and began the PP control condition and the students in the PP control condition crossed over and began the RMF treatment condition for the last six weeks of the study. See Table 7 for a layout of the design.

Table 7

Study Design

Factors	Pre-test		Intervention 1 (6 weeks)	Test1				Intervention 2 (6 weeks)	Test2			
Test	easy	AIMS		Easy		AIMS			easy		AIMS	
Grade	On	On		On	Off	On	Off		On	Off	On	Off
Treatment	X	X	Computer	X	X	X	X	Traditional	X	X	X	X
Control	X	X	Traditional	X	X	X	X	Computer	X	X	X	X

Note. easy = easyCBM assessments; AIMS = AIMSweb assessments; On = grade-level test; Off = subsequent-grade-level test.

Procedures. I created a website with a calendar to keep track of the important dates for the study, to make classroom practice and assessment instructions easily available to the teachers, and to allow teachers to blog about any difficulties that occurred throughout the study, such as issues with internet access. I also uploaded all three classes to the easyCBM website and designated the first of the tests that students would take at each grade level. I chose the first of the progress monitoring tests in each of the three focal points. I met with the three teachers before students completed the pre-test to show them how to navigate the website and answered questions they had about their responsibilities.

The teachers were already familiar with the AIMSweb test administration and scoring materials from the school's adoption of the assessment in 2008, so we reviewed the information and ensured that they had the materials they needed for the three testing periods. I showed teachers how the students would log-in to easyCBM, and we bookmarked the website on the computers so that students could access it easily during testing times. I gave each teacher a crate with file folders for their grade level of the math worksheets, answer keys, and packets of flash cards for students in the PP control condition, telling them to instruct students to work by themselves when working with the flashcards and to explain the worksheet testing procedure to the students. Finally, we reviewed the steps for showing students how to log-in to the *MathFacts in a Flash* program to practice and test so that they could teach the RMF groups these steps.

All students completed the first easyCBM and AIMSweb progress monitoring measures at their grade level to determine groups and create a baseline. In each grade level, the RMF group was taught by their teacher to use *MathFacts in a Flash* for math facts practice and testing. The PP group was instructed by their teacher to work independently to practice math facts with the flash cards, and the group was informed of the procedure for testing with the worksheets. Both groups followed the same scope and sequence of math facts practice and testing. At the start of the math period in each grade, the RMF groups used MacBook Air laptops to log-in to *MathFacts in a Flash* for 10 minutes every day over the course of a week, while the PP groups practiced their math facts using flash cards. See Table 8 for the essential practice and testing differences between the RMF and PP groups.

Table 8

Essential Practice and Testing Differences Among the RMF and PP Groups

Quality	RMF	PP
Mode of delivery	MacBook Air laptops	Flash cards and worksheets
Interactive features	Automatic feedback on incorrect answers with correct answer highlighted. Required to answer a missed problem immediately after seeing the correct answer. Immediate fluency and accuracy scores upon completion of session with incorrect problems displayed.	Front and back flash cards show the answer on the back of the card for immediate feedback. Students wait until the end of the day to get their test back with the incorrect answers marked and their accuracy score.
Personalization	Students practice missed facts from last test or practice session with other frequently missed and known facts at their level. Students may test at any time to advance to next level.	Students at the same level receive the same flash cards and worksheets. Students must wait until two minutes before the end of the practice session to be timed for testing.
Goal Setting	100% accuracy and 2 min. time goal. Students who completed all levels set time goals to decrease to 1 min.	100% accuracy and 2 min. time goal. Students who completed all levels set time goals to decrease to 1 min.

Note. RMF= treatment condition; PP = control condition.

Toward the end of the second week of the study, I noticed that several students were rapidly completing the levels on the computer for their grade. I decided to allow those students who met the 100% accuracy and two minute time goal for all of the levels at their grade to advance to the next grade level fact scope and sequence for both the

RMF and PP group. Students who completed the next grade level fact sequence returned to the beginning of their own grade level facts and retested for a faster goal time decreasing in 15-second intervals to one minute.

Two days before the end of six weeks, all students completed the second progress monitoring easyCBM and AIMSweb math tests for their respective grade levels. The last day of the sixth week, all students completed the next grade level's progress monitoring easyCBM and AIMSweb math tests. Then, students in the RMF groups were instructed to practice their math facts with flash cards and paper based math sheets, and the PP groups were taught to use *MathFacts in a Flash* to practice and test math facts. The two crossed groups continued the scope and sequence of facts from where they stopped in the alternate treatment so that they continued to practice with content ordered in a logical sequence. Two days before the twelfth and final week, all students completed the third easyCBM and AIMSweb progress monitoring math tests for their respective grade levels. The last day of the study, all students completed an alternate form of the next grade level's progress monitoring easyCBM and AIMSweb math test. Table 9 displays an overview of the crossed design for both groups.

Table 9

Overview of the Crossed Design for the RMF and PP Groups

Timeline	RMF	PP
Prior to study	Complete first easyCBM and AIMSweb progress monitoring tests and learn to use *MathFacts in a Flash* software	Complete first easyCBM and AIMSweb progress monitoring tests and instructed to use flash cards and worksheets for practice

Table 9 (continued)

Timeline	RMF	PP
Weeks 1-5	Practice *MathFacts in a Flash* software daily for 10 minutes and when ready to advance to the next facts, take a timed test on the computer	Practice with flash cards daily for 10 minutes and when ready to advance to the next facts, take a timed test from a worksheet two minutes before the end of session
Week 6	Practice as before, complete the second easyCBM and AIMSweb tests 2 days before week's end, complete next grade level easyCBM and AIMSweb tests at week's end, and learn to use flash cards and worksheets for practice	Practice as before, complete the second easyCBM and AIMSweb tests 2 days before week's end, complete next grade level easyCBM and AIMSweb tests at week's end, and learn to use *MathFacts in a Flash* software
Weeks 7-11	See weeks 1-5 for PP Group	See weeks 1-5 for RMF group
Week 12	Practice as before, complete the third easyCBM and AIMSweb progress monitoring math tests at and above grade level	Practice as before, complete the third easyCBM and AIMSweb progress monitoring math tests at and above grade level

Note. RMF = treatment condition; PP = control condition.

Data Analysis

I rank-ordered students by their easyCBM pre-test scores and then randomly assigned them to either the control or treatment condition to ensure that the two groups were equal. The two students with the highest scores were randomly assigned to either the control or treatment condition, followed by a random assignment of the next two highest scoring students until all students were assigned in each grade. Equivalent groups were essential to my study because students were not randomly selected to participate and because they would be receiving the treatment and control conditions at different times in the study.

I then grouped the raw scores for the two conditions (RMF treatment and PP control) at each grade level for each assessment (easyCBM and AIMSweb) to create the maximum number of students per condition in each grade. I conducted a correlated-groups *t*-test to compare the two conditions (RMF and PP) on the AIMSweb on-grade-level measures and easyCBM on-grade-level measures for each grade. I conducted a supplementary analysis on the AIMSweb and easyCBM off-grade-level measures as well. I used an independent-groups *t*-test to compare the two conditions (RMF and PP) on the AIMSweb and easyCBM off-grade-level measures for each grade.

CHAPTER III

RESULTS

In this section, I report the results of the study and describe the strategies for analyzing the easyCBM and AIMSweb data. Results from these data were used to answer the primary research question: Do third, fourth, and fifth-grade students who use *MathFacts in a Flash* software to practice basic math facts score significantly higher on mathematics curriculum-based measures than third, fourth, and fifth-grade students who use flashcards and worksheets for practice?

Descriptive Statistics

Students completed alternate forms of the standardized easyCBM and AIMSweb assessments at the beginning, middle, and end of the study for their respective grades. I chose easyCBM because it was designed to focus on student conceptual understanding of the National Council of Teachers of Mathematics Focal Point Standards (Numbers and Operations, Measurement and Data Analysis, and Numbers and Operations and Algebra) rather than computational skills (Alonzo & Tindal, 2010). I chose AIMSweb M-COMP to focus specifically on computational skills (Pearson Education, 2010). I also chose easyCBM and AIMSweb because of the moderate to high reliability of the alternate forms for each measure previously mentioned in the measurements section of Chapter II. Reliable alternate forms were important to this study because students received the treatment condition at different times, and I needed to be sure that the tests they took were of equivalent difficulty.

Additionally, students completed alternate forms for the subsequent grade level easyCBM and AIMSweb assessments after the first six weeks of the study and after the last six weeks of the study. Students completed these additional assessments at the next grade level because they were allowed to advance to the next grade level's math fact scope and sequence when they completed the scope and sequence at their own grade level. The reasoning behind this approach was that if math fact practice has an effect on math scores at grade level, then math fact practice at the next grade level might have an effect on math scores for the next grade level.

The number of students in each grade who advanced to the next-grade-level scope and sequence and the number of students who re-tested for faster time goals are displayed in Table 10. The scope and sequence for *MathFacts in a Flash* ends at grade 5; therefore, the fifth-graders were the only students to first retest for a faster time goal at their own grade level before moving on to a more difficult sequence of practice.

Table 10

Student Progression Through Scope and Sequence and Time Goals

	Grade 3 (*N* = 20)		Grade 4 (*N* = 20)		Grade 5 (*N* = 18)	
Weeks / Progress	RMF	PP	RMF	PP	RMF	PP
1-6						
Next grade	6	3	6	3		
Faster time goals					6	2
7-12						
Next grade	2	5	4	5	2	3
Faster time goals	4	4	4	4	4	4

Note. RMF = treatment condition; PP = control condition.

The raw score descriptive statistics for each grade are displayed in Tables 11 through 13. Each group (RMF and PP) contains different students at different testing points in the study. Raw score descriptive statistics disaggregated by gender and ethnicity are available in Appendix D.

Table 11

Third Grade Raw Score Descriptive Statistics by Groups and Measures in the Crossed Design

	Total (*N* = 20)		RMF (*n* = 10)		PP (*n* = 10)	
Test	*M*	*SD*	*M*	*SD*	*M*	*SD*
easyCBM						
Pretest	35.80	3.25	35.80	3.46	35.80	3.23
On-grade-level post 1	36.80	3.67	36.10	3.41	37.50	3.95
On-grade-level post 2	41.15	3.30	42.20	3.26	40.10	3.14
Off-grade-level pretest	32.80	2.90	32.70	3.43	32.90	2.47
Off-grade-level post	32.65	3.08	33.20	2.74	32.10	3.45
AIMSweb						
Pretest	45.60	11.20	45.40	13.24	45.80	9.46
On-grade-level post 1	52.10	11.75	51.80	13.72	52.40	10.16
On-grade-level post 2	56.50	9.60	58.30	10.89	54.70	8.33
Off-grade-level pretest	29.40	8.94	27.70	7.85	31.10	10.04
Off-grade-level post	34.10	6.55	33.10	7.17	35.10	6.08

Note. Shaded areas between the two groups show the places at which the students crossed to the opposing treatment condition. RMF = treatment condition; PP = control condition; Off-grade-level = assessments at the next grade level.

Table 12

Fourth Grade Raw Score Descriptive Statistics by Groups and Measures in the Crossed Design

	Total (*N* = 20)		RMF (*n* = 10)		PP (*n* = 10)	
Test	*M*	*SD*	*M*	*SD*	*M*	*SD*
easyCBM						
Pretest	34.70	4.84	34.40	5.70	35.00	4.08
On-grade-level post 1	35.10	5.83	34.70	7.45	35.50	3.98
On-grade-level post 2	33.80	5.24	33.80	4.16	33.80	6.37
Off-grade-level pretest	28.10	5.93	29.90	5.53	26.30	6.04
Off-grade-level post	30.00	6.04	29.80	6.51	30.20	5.87
AIMSweb						
Pretest	38.50	13.62	37.70	16.36	39.30	11.08
On-grade-level post 1	44.20	13.72	42.30	17.43	46.10	9.24
On-grade-level post 2	39.90	14.90	43.00	13.03	36.80	16.66
Off-grade-level pretest	14.80	8.94	15.10	9.99	14.50	8.28
Off-grade-level post	13.85	7.01	14.70	7.45	13.00	6.83

Note. Shaded areas between the two groups show the places at which the students crossed to the opposing treatment condition. RMF = treatment condition; PP = control condition; Off-grade-level = assessments at the next grade level.

Table 13

Fifth Grade Raw Score Descriptive Statistics by Groups and Measures in the Crossed Design

	Total (N = 18)		RMF (n = 9)		PP (n = 9)	
Test	*M*	*SD*	*M*	*SD*	*M*	*SD*
easyCBM						
Pretest	32.50	6.95	31.89	8.02	33.11	6.11
On-grade-level post 1	35.44	6.96	35.11	7.74	35.78	6.55
On-grade-level post 2	37.89	7.37	38.22	7.23	37.56	7.94
Off-grade-level pretest	28.28	6.46	27.56	4.16	29.00	8.38
Off-grade-level post	28.17	6.29	29.22	5.07	27.11	7.47
AIMSweb						
Pretest	30.22	16.44	30.78	20.25	29.67	12.81
On-grade-level post 1	32.72	16.98	32.56	19.66	32.89	15.05
On-grade-level post 2	37.44	17.87	37.22	14.94	37.67	21.34
Off-grade-level pretest	22.06	11.84	20.44	8.69	23.67	14.71
Off-grade-level post	29.17	12.88	27.78	11.60	30.56	14.61

Note. Shaded areas between the two groups show the places at which the students crossed to the opposing treatment condition. RMF = treatment condition; PP = control condition; Off-grade-level = assessments at the next grade level.

Comparison of Treatments by Grade

To analyze the effects of the computer software program at each grade level, the raw scores from students in the RMF treatment in the last six weeks of the study were grouped with the raw scores from students in the RMF treatment in the first six weeks of the study (e.g., all third grade raw scores in the RMF treatment were grouped, regardless of the order in which the treatment was received). Thus, the number of students in each condition was maximized, which yielded greater statistical power. The same is true for the data on the students in the PP control condition. This was possible due to the equivalent forms taken by students throughout the study for both on-grade AIMSweb and easCBM measures and the random assignment of students to equivalent groups at the beginning of the study.

On-grade-level measures. I conducted a correlated-groups *t*-test to compare students' on-grade-level easyCBM and AIMSweb scores in the RMF treatment and PP control conditions for each grade level using SPSS 17.0. I used an alpha level of .05 for each test. I chose this analytic technique because of the within participants design (each student had a score for both treatment and control conditions); and because students had been randomized to each condition, the groups were roughly normally distributed and homogenous (Jackson, 2010).

The results of the correlated-groups *t*-tests for each grade indicated that there was not a significant difference in the scores for the RMF treatment and PP control conditions for either the easyCBM or the AIMSweb on-grade-level measures. Table 14 displays the results of the correlated-groups *t*-test for each condition by grade. These results suggest that the RMF treatment did not have a greater effect on student achievement than the PP control at any grade. Specifically, the results suggest that when students practiced math facts with *MathFacts in a Flash*, they did not perform differently than when they practiced with paper-pencil methods.

Table 14

Correlated-Groups t-Test Results of the On-Grade-Level Measures for Each Grade

	RMF		PP				95% CI	
Test/Grade	*M*	*SD*	*M*	*SD*	*t*(19)	*p*	*LL*	*UL*
easyCBM/								
3	39.15	4.51	38.80	3.72	.28	.78	-2.25	2.95
4	34.25	5.89	34.65	5.24	-.50	.62	-2.08	1.28
5	36.67	7.44	36.67	7.12	.00*	1.00	-2.96	2.96
AIMSweb/								
3	55.05	12.51	53.55	9.12	.68	.50	-3.11	6.11
4	42.65	14.98	41.45	13.95	.54	.60	-3.49	5.89
5	34.89	17.11	35.28	18.08	-.17*	.87	-5.15	4.38

Note. CI = confidence interval; *LL* = lower limit; *UL* = upper limit; * = *t*(17).

Off-grade-level measures. I conducted a supplementary analysis for the off-grade-level measures. Because students did not take the off-grade-level measures until the middle and end of the study, the post-test scores represent a smaller sample. In essence, the sample size for each condition was half of the sample size for the on-grade-level measures.

I conducted an independent-groups *t*-test to compare students' off-grade-level easyCBM and AIMSweb scores in the RMF treatment and PP control conditions for each grade level using SPSS 17.0. I used an alpha level of .05 for each test. I chose this analytic technique because of the between-participants design at this point of the study (each student had only one score for either the treatment or the control condition); and because students had been randomized to each condition at the start of the study, the groups were roughly normally distributed and homogenous (Jackson, 2010).

The results of the independent-groups *t*-tests for each grade indicated that there was not a significant difference in the scores for the RMF treatment and PP control conditions for either the easyCBM or the AIMSweb off-grade-level measures. Table 15 displays the results of the independent-groups *t*-test for each condition by grade. Similar to the results for the on-grade-level measures, these results suggest that the RMF treatment did not have a greater effect on student achievement than the PP control at any grade. Specifically, the results suggest that when students practiced math facts with *MathFacts in a Flash* at the next grade level, they did not perform differently than when they practiced with paper-pencil methods at the next grade level.

Table 15

Independent-Groups t-Test Results of the Next-Grade-Level Measures for Each Grade

	RMF		PP				95% CI	
Test/Grade	*M*	*SD*	*M*	*SD*	*t*(18)	*p*	*LL*	*UL*
easyCBM/								
3	33.20	2.74	32.10	3.45	-.79	.44	-4.03	1.83
4	29.80	6.51	30.20	5.87	.14	.89	-5.42	6.22
5	29.22	5.07	27.11	7.47	-.70*	.49	-8.49	4.27
AIMSweb/								
3	33.10	7.17	35.10	6.08	.67	.51	-4.25	8.25
4	14.70	7.45	13.00	6.83	-.53	.60	-8.42	5.02
5	27.78	11.60	30.56	14.61	.45*	.66	-10.40	15.96

Note. CI = confidence interval; *LL* = lower limit; *UL* = upper limit; * = *t*(16).

CHAPTER IV

DISCUSSION

The results of the study indicate that for on-grade-level curriculum-based measures (easyCBM and AIMSweb), there was not a significant difference in math scores between students practicing math facts with *MathFacts in a Flash* and those practicing math facts with flash cards and worksheets. In addition, the results for the off-grade-level curriculum-based measures (easyCBM and AIMSweb) indicated no significant differences in math scores between students practicing with *MathFacts in a Flash* and those practicing with flash cards and worksheets. In essence, the results suggest that there is no difference between teacher-created and computer-driven interventions when specific qualities of teaching and learning are present in both conditions. In the following sections I will interpret the results of the study in the context of the literature, explain the threats to validity that existed in my study, and propose implications for leaders in education.

Interpretations

I designed the study to take place over the course of 12 weeks because previous research has indicated that interventions have been found to have greater impact on student achievement when they are less than a year (Li & Ma, 2010). Alternately, a meta-analysis by Kulik and Kulik (1991) found a larger effect size for interventions that lasted four weeks or less, but the studies included in that analysis were not specifically focused on computer-based math interventions. Students who participated in this study were engaged in math fact practice and testing for ten minutes every day, which was in the 3-5 days a week for 5-15 minutes a day range recommended by the *MathFacts in a*

Flash software program (Ysseldyke et al., 2005) and corresponds to research-based recommendations for math fact practice (Gersten et al., 2009; Kulik, 2003). Although students tested for a new level during the ten-minute session, they still received approximately eight minutes of practice per session.

Students were able to set time goals to work toward mastery (Blanton et al., 1997; Irish, 2002; Martindale et al., 2005; Wittman et al., 1998; Xin, 1999), but they were not allowed to do this until they had completed the pre-determined two-minute time goal for all levels at their grade and the next grade. Nor were they able to set accuracy goals. Perhaps it would have made a positively significant difference in math scores had they been able to set their own time goal at the start of the study and a fluency goal on their own (Blanton et al., 1997; Irish, 2002; Martindale et al., 2005; Wittman et al., 1998) based on their own sense of readiness (Tomlinson, 1999).

The *MathFacts in a Flash* program differentiated content for students using compacting (Tomlinson, 1999) in which each student began the program with a pretest that determined what the student knew, what challenges the student faced, and then moved the student into targeted practice or on to a new level. The only information I was able to obtain about how the program targeted practice to an individual student was the information from the company about repeated practice on unknown facts and other commonly-missed facts for each level (Renaissance, 2009). I was able to use the program to determine that for each 20-question practice session, any missed facts from my previous session were repeated on average two times. Perhaps this was not enough targeted practice for students practicing with computer software to result in a significantly positive difference in math scores.

The *MathFacts in a Flash* program provided students with differentiation in environment (Tomlinson, 1999) through the visual feedback they received for incorrect answers. Students simultaneously saw a red X on the incorrect choices they made and a red box around the correct answer, which set them up for correctly answering the same question directly after being prompted. The program did not provide any other animated graphics (Martindale et al., 2005; Xin, 199) or sounds (Blanton et al., 1997; Irish, 2002; Xin, 1999). There might have been a positive increase in math scores for the students practicing with *MathFacts in a Flash* had animation and sound been increased or present in the program.

Students were assessed with criterion-referenced curriculum-based measures (Blanton et al., 1997; Martindale et al., 2005; Quinn & Quinn, 2001; Royer et al., 1994; Salerno, 1995), which have been shown to be reliable measures of computational knowledge (Thurber et al., 2002). In their meta-analysis, Li and Ma (2010) found a larger effect size (ES = 1. 02) for studies that used teacher or researcher-made tests to determine the effectiveness of the treatment or intervention. I did not proceed with a teacher-made test, because the easyCBM and AIMSweb assessments were valid, reliable tests of student conceptual knowledge of math standards and accuracy and fluency with math computation, respectively (Nese et al., 2010; Pearson Education, 2010). Additionally, Li and Ma (2010) found that standardized tests (tests with good psychometric properties), such as the assessments I chose, had a moderate effect size (ES = 0.78), which I was comfortable accepting over the effect size possible with teacher- or researcher-made tests.

Although students completed multiple tests throughout the study, the design was essentially a pre-test post-test design. In the meta-analysis by Li and Ma (2010), seven of the eight studies that I focused on referencing third, fourth, and/or fifth-graders used a pre-test post-test design (Blanton et al., 1997; Martindale et al., 2005; Quinn & Quinn, 2001; Royer et al., 1994; Salerno, 1995; Wittman et al., 1998; Xin, 1999). Even though I wanted to conduct a study using a multiple baseline design in order to see the intricacies of any changes over time (Babbie, 2010), the option I chose for trying to increase a small sample size (the crossed design) and my method for analyzing the data (*t*-tests at each grade for each measure) required a pre-test post-test design.

Finally, although I was unable to detect any significant differences between the two groups in each grade, I was able to make comparisons between participants' scores and the norm groups' scores in both the easyCBM and AIMSweb systems. According to the *easyCBM Progress Monitoring Score Interpretation Guidelines* (2011), third grade students in the RMF treatment condition began the study at the 50th percentile for the winter benchmark and completed the study at the 75th percentile for the spring benchmark, whereas the third grade students in the PP control condition began and ended the study in the 50th percentile. In contrast, the fifth grade students in both conditions showed the same growth from approximately the 20th percentile to approximately the 50th percentile. The easyCBM fourth grade norm group outperformed the fourth grade students in both treatment conditions. Fourth-grade students in my sample showed no increase in performance when compared to normative performance, but rather a decrease from winter to spring. It is possible that the norm comparisons at grade level might have been influenced by student advancement to the subsequent grade level as seen in Table

10 from Chapter III. Thus, although there was no significant difference in performance between the treatment and control groups in my study, third and fifth grade students in both conditions made more progress than the norm groups at their grade levels, and students in the fourth-grade sample slightly underperformed, when compared to expected normative growth from winter to spring. This finding supports the assertion that the structured math skills practice used in both the treatment and control conditions was effective in helping students develop their math conceptual understanding.

Students' growth in computation skills was comparable to the observations of growth I made regarding the easyCBM system. According to the *AIMSweb National Norms Table* (2012), the third graders in the RMF treatment scored in the 50^{th} percentile in the winter and a few points above the 50^{th} percentile in the spring, whereas the third graders in the PP control began and ended the study in approximately the 50^{th} percentile. The fourth graders in the RMF treatment maintained growth in the 25^{th} percentile, whereas the fourth graders in the PP control scored below the 25^{th} percentile in the spring. The fifth graders in both conditions performed similarly with growth from the 50^{th} percentile to approximately the 75^{th} percentile. Again, despite there being no statistically significant difference between the two conditions, it is worth pointing out that students in my study made progress in their math computation skills that was greater than or at least the same as the progress made by the norm group in the AIMSweb system.

Limitations

One threat to the external validity of this study is sampling (Trochim, 2006) as students were not randomly selected but were included as part of a convenience sample. The best way for me to control for this threat is to describe them (Babbie, 2010). They are a selection of students from an archdiocesan population that is similar to the population of Catholic schools across the country (McDonald & Schultz, 2010), but any generalizability to populations outside of the Catholic schools should be made with caution.

One threat to internal validity is in regards to history (Campbell & Stanley, 1963). During the course of the study, two events occurred that may have had an effect on the results of the study. One that I think might be significant is the timing of the study. Because I was unable to begin the study at the start of the school year, I began the study after the winter break, which means that students were off school for an entire week for the spring break vacation during the study. Although this time off was not for multiple weeks, there may have been a decline in student attitude and focus prior to and after the break (Fraenkel & Wallen, 2003).

The second event that occurred was a change in teachers for the fourth grade at the start of the fifth week of the study because of a maternity leave. This change in teachers was a threat to internal validity because the long-term substitute had less teaching experience than the classroom teacher and because of a possible change in consistency and expectations with a new teacher. Even though the new teacher received the same amount of training and assistance related to this specific study as the other teachers, her arrival mid-study may have introduced variability in the delivery of the intervention and administration of the measures. The students knew from the start of the year that their teacher would be going on maternity leave and seemed to do well with the substitute, but they may have performed differently for her than they would have for their regular classroom teacher.

An additional threat to the internal validity of the study relates to fidelity of implementation. I was unable to observe during the practice and testing sessions to be sure that the procedures were consistent, but relied on the daily posts to the blog about problems that arose, the tracking of practice and testing for both groups (control and treatment), and conversations with the teachers on breaks or after school.

A final threat to the study involves instrumentation (Campbell & Stanley, 1963). In this study, the availability of technology was extremely important to the outcomes. At the end of the third week and through most of the fourth week of the study, Internet connectivity was a challenge for teachers and students. The school had purchased routers that would not allow for the high traffic needed to connect the administration, teachers, and students to the Internet. It took a week for the issue to be resolved. I had not taken into account the need for an alternative option and set of procedures when Internet or hardware failed before the study began. I created a protocol to use math apps on the iPads, but before that, teacher problem-solving related to how they dealt with the lack of connectivity and subsequent inability to use the online math practice program varied.

After the week of trouble with Internet access, there were two days in which one class was not able to practice on the computers because the charging cart had been damaged and the computers did not have enough charge to sustain student use. This was two days before spring break. The cart was fixed and computers charged when students returned. The issues with Internet connectivity and hardware may have negatively impacted the results of the study because students in the computer intervention did not have access to the *MathFacts in a Flash* treatment at various times in the study.

Implications

Even though there are many possible reasons for the outcomes in my study, it is unclear why students in the RMF treatment condition did not perform significantly better than students in the PP control condition; however, they also did not perform significantly worse. Given these findings, I will suggest some implications for school leaders in considering the adoption of a personalized software program as opposed to implementing traditional paper pencil practice for learning math standards.

The first implication is that the infrastructure beneath any form of technological adoption must be able to support the traffic of its users. Adopting a computer program that is inaccessible renders the adoption ineffective. Building a strong enough infrastructure and purchasing the hardware and equipment necessary to carry out the adoption of a technology program can also be costly. In addition, possible glitches in Internet connectivity and hardware issues can cause trouble. This may be a reason why 61% of fourth-grade students report that they *never or hardly ever* use computers for math at school (NAEP, 2009).

The second implication is that previous research findings have shown that there is a moderate, but significantly positive effect of computer technology on mathematics achievement (Li & Ma, 2010). Given that differences in gain scores for the RMF treatment group and the PP control group were not significantly different, leaders may still consider the *MathFacts in a Flash* software as a worthwhile adoption for various other reasons supported by research. The *MathFacts in a Flash* program allows students to set learning goals to work toward mastery (Blanton et al., 1997; Irish, 2002; Martindale et al., 2005; Wittman et al., 1998; Xin, 1999). As previously mentioned, I did not allow students to set their own fluency and accuracy goals, but the program does allow them to do this in collaboration with their classroom teacher. The program provides students with feedback on their progress toward their learning goals (Irish, 2002; Martindale et al., 2005; Xin, 1999). They can see the errors they make during practice sessions when they make them and they receive a list of their errors along with their fluency and accuracy score that they can print at the end of their session.

In addition to goal setting and continuous feedback, the program provides for personalization or differentiation through the use of compacting (Tomlinson, 1999). Students progress through the program based on their achievement on a pre-test. Each student is able to work at an individualized pace and level. Parents can opt to receive e-mail notifications generated by the program when their child completes a level, and students can access their accounts at home as well. Although students are not able to test into a new level, they do have a secure login that provides the option to practice from home. The study did not include the use of these options; however, the additional practice features outside of the classroom may strengthen student fluency and accuracy overall. Increased fluency and accuracy with math facts may lead to increased math standards achievement for students.

A third implication for leaders is in the importance of training and consistency. The change in teachers for the fourth grade may have had an effect on how students at that grade level performed. I probably should have made arrangements to observe the practice sessions to ensure that the interim teacher and probably all teachers were following the 10-minute practice protocol. I draw attention to this implication because I feel that it demonstrates the importance of being present in the classroom to guide and provide teachers with support. Not to punish, but to help increase the likelihood that consistency in teaching and learning will have positive effects on student achievement in all subjects.

Although I was unable to detect a significant difference between *MathFacts in a Flash* and paper-pencil conditions, an underlying implication is that even though there was a delay in feedback to students and a lack of computer animation in the paper-pencil condition, the computer software program was not any more effective than the teacher-created program. So why implement a program like *MathFacts in a Flash*?

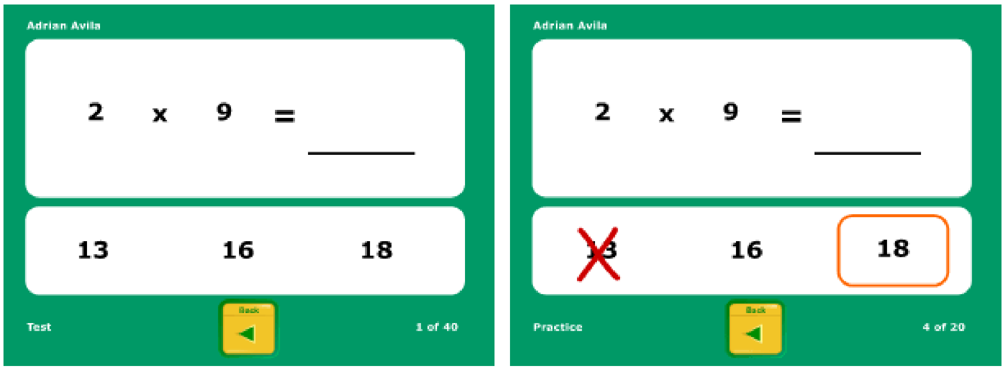
Programs like *MathFacts in a Flash* can assess, monitor, and challenge students while also having small positive effects on students' attitudes toward instruction, coursework, and computers (Kulik & Kulik, 1991). A benefit to teachers is that such programs require less grading and paperwork maintenance than typical paper-pencil methods. These programs do not produce significant negative effects on student achievement. In fact, my study suggests that differentiated computer software is just as effective at increasing student achievement in math standards as highly-structured paper-pencil methods. In other words, differentiated software is an acceptable substitute for the teacher, for practicing math facts. The implications of replacing teachers with differentiated software then turn toward more research.

Programs like *MathFacts in a Flash* have a place in the math classroom because they allow the teacher to spend more time with students while best practices in differentiation are being implemented through the use of computer software. Implications from this study suggest a need for further research in education regarding the instructional practices of teachers. If tutorial software is implemented to not only improve student achievement, but also to provide more time for the teacher, how can the teacher more effectively use the time that would have been spent on grading, organizing, testing, and providing feedback to students to significantly affect student achievement in math standards?

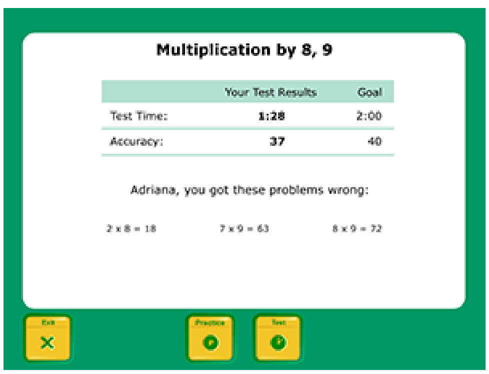
Finally, further research is needed to more fully investigate the specific qualities of personalization or differentiation in tutorial software. There is a lack of research available to help educators identify the factors that impact the effectiveness of instructional software in mathematics (NMAP, 2008). Additional research on the important factors of instructional software can help schools determine how it can better influence student math-standards achievement.

APPENDIX A

MATHFACTS IN A FLASH PROGRAM



Screen shots for *MathFacts in a Flash* software



Student report for *MathFacts in a Flash* software

Table 16

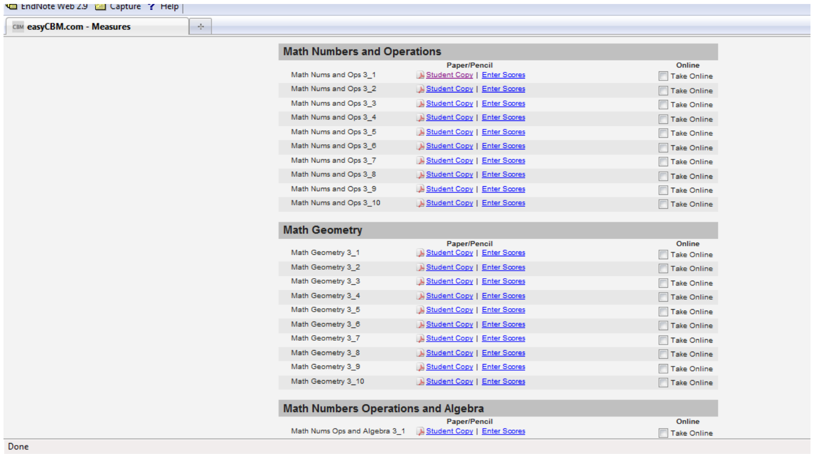
Scope and Sequence by Grade

Grade 3 (Multiplication)	Grade 4 (Division)	Grade 5 (Advanced)
0,1	*Review: +, -, x	*Mixed Review +, -, x, /
2,3	1, 2	Squares to 15, 20
4, 5	3, 4	Squares Review
0 to 5	5, 6	*Review: +, -, x, /, squares
6, 7	1 to 6	Fractions to Decimals
8, 9	7, 8	Decimals to Fractions
10	9, 10	Percentages to Decimals
6 to 10	7 to 10	Decimals to Percentages
11, 12	11, 12	Fractions to Percentages
*Review 1	*Review 1	Conversion Review
*Review 2	*Review 2	* +, -, x, /, squares, conversion
*Review: +, -, x	*Mixed Review +, -, x, /	

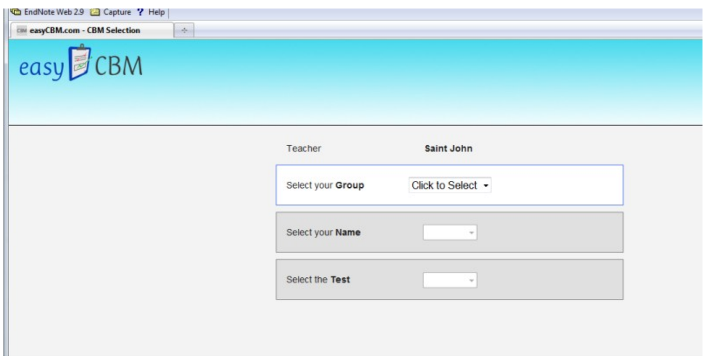
Note. * = Multiplication and division reviews do not include items containing the numbers 11 and 12.

APPENDIX B

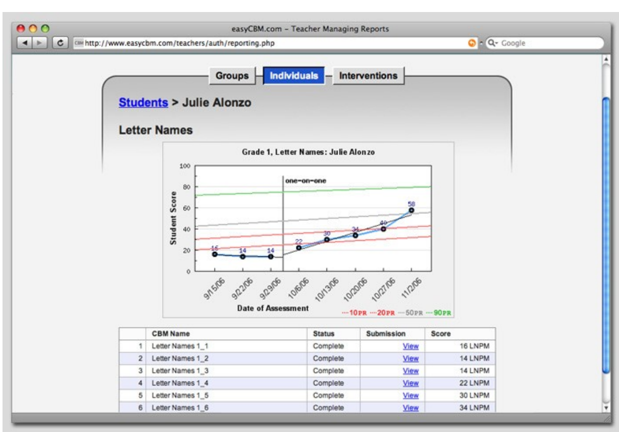
EASYCBM SAMPLES



easyCBM math tests for the three focal points that students take on the computer



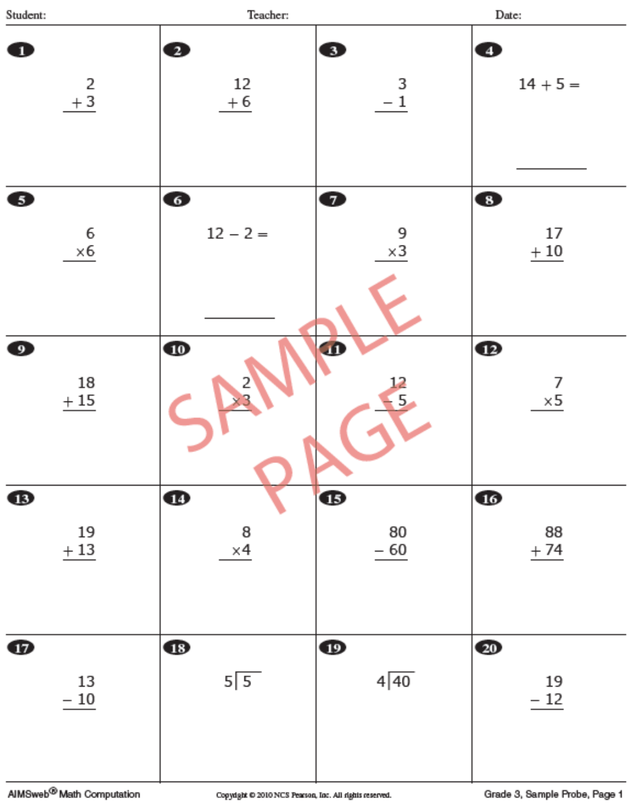
easyCBM student login screen shot



easyCBM example report

APPENDIX C

AIMSWEB SAMPLES


Student: Teacher: Date:

1. 2 + 3
2. 12 + 6
3. 3 − 1
4. 14 + 5 = ____
5. 6 × 6
6. 12 − 2 = ____
7. 9 × 3
8. 17 + 10
9. 18 + 15
10. 2 × 3
11. 12 − 5
12. 7 × 5
13. 19 + 13
14. 8 × 4
15. 80 − 60
16. 88 + 74
17. 13 − 10
18. 5)5
19. 4)40
20. 19 − 12

AIMSweb® Math Computation Grade 3, Sample Probe, Page 1

Example M-COMP Probe.

Grade 5, Probe 3 Answer Key

Item No.	Answer	Correct	Incorrect	Item No.	Answer	Correct	Incorrect
1.	128	1	0	21.	79	1	0
2.	4	1	0	22.	$\frac{63}{100}$	2	0
3.	$\frac{1}{8}$	2	0	23.	16 r3; 16.6; 16 $\frac{3}{5}$	3	0
4.	143	2	0	24.	5	3	0
5.	820	1	0	25.	396	1	0
6.	$\frac{1}{12}$	3	0	26.	$\frac{9}{13}$	1	0
7.	70	1	0	27.	$\frac{21}{31}$	2	0
8.	0.95	1	0	28.	45	3	0
9.	19	2	0	29.	$\frac{1}{12}$	1	0
10.	$\frac{1}{2}$	2	0	30.	$\frac{1}{15}$	3	0
11.	10 r3; 10.6; 10 $\frac{3}{5}$	2	0	31.	53 r3; 53.5; 53 $\frac{3}{6}$; 53 $\frac{1}{2}$	3	0
12.	3883	2	0	32.	33	3	0

Example Answer Key M-COMP

APPENDIX D

DISAGGREGATED RAW SCORES

Table 18

Fourth Grade Raw Scores by Gender and Ethnicity

	Gender		Ethnicity
	Male (N = 9)	Female (N = 11)	Caucasian (N = 18)
Test	*M (SD)*	*M (SD)*	*M (SD)*
easyCBM			
Pretest	34.44 (5.64)	34.91 (4.35)	35.22 (4.71)
On-grade-level posttest 1	36.33 (4.12)	34.09 (6.96)	35.39 (6.07)
On-grade-level posttest 2	35.67 (5.52)	32.27 (4.69)	34.11 (5.44)
Off-grade-level posttest 1	29.44 (4.85)	27.00 (6.71)	28.56 (6.00)
Off-grade-level posttest 2	31.00 (5.77)	29.18 (6.40)	30.06 (6.26)
AIMSweb			
Pretest	39.44 (14.91)	37.73 (13.16)	39.50 (13.56)
On-grade-level posttest 1	45.22 (11.34)	43.36 (15.90)	45.22 (13.85)
On-grade-level posttest 2	40.56 (14.45)	39.36 (15.94)	41.17 (14.37)
Off-grade-level posttest 1	15.56 (8.46)	14.18 (9.67)	14.50 (8.68)
Off-grade-level posttest 2	14.22 (5.36)	13.55 (8.38)	13.94 (7.28)

Note. The scores for 1 student of "Other" ethnicity and for 1 student of "Asian" ethnicity are not reported in this table. Off-grade-level = assessments at the next grade level.

Table 19

Fifth Grade Raw Scores by Gender and Ethnicity

	Gender		Ethnicity	
	Male (*N* = 7)	Female (*N* = 11)	Caucasian (*N* = 15)	Asian (*N* = 2)
Test	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
easyCBM				
Pretest	29.00 (6.83)	34.73 (6.33)	32.53 (6.69)	38.00 (1.41)
On-grade-level posttest 1	31.57 (6.19)	37.91 (6.50)	35.07 (5.71)	45.00 (1.41)
On-grade-level posttest 2	35.43 (7.39)	39.45 (7.26)	38.27 (6.40)	43.50 (.71)
Off-grade-level posttest 1	25.43 (4.89)	30.09 (6.88)	27.27 (4.99)	40.00 (1.41)
Off-grade-level posttest 2	25.57 (6.16)	29.82 (6.06)	27.80 (5.23)	36.50 (4.95)
AIMSweb				
Pretest	18.86 (7.95)	37.45 (16.55)	27.67 (12.09)	61.00 (4.24)
On-grade-level posttest 1	23.14 (13.06)	38.82 (16.82)	30.60 (11.76)	*
On-grade-level posttest 2	28.43 (16.50)	43.18 (16.92)	35.93 (14.73)	64.00 (4.24)
Off-grade-level posttest 1	14.43 (7.64)	26.91 (11.68)	20.33 (7.51)	45.50 (2.12)
Off-grade-level posttest 2	22.86 (12.05)	33.18 (12.21)	28.20 (8.70)	50.50 (6.36)

Note. The score for 1 student of "Other" ethnicity is not reported in this table. Off-grade-level = assessments at the next grade level; * = student scores are the same.

Table 17

Third Grade Raw Scores by Gender and Ethnicity

	Gender		Ethnicity		
	Male (*N* = 7)	Female (*N* = 13)	Caucasian (*N* = 15)	Hispanic (*N* = 2)	Asian (*N* = 2)
Test	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
easyCBM					
Pretest	36.29 (2.50)	35.54 (3.67)	35.27 (3.13)	38.50 (4.95)	37.00 (4.24)
On-grade-level posttest 1	37.71 (2.29)	36.31 (4.23)	36.33 (3.69)	38.50 (4.95)	38.50 (4.95)
On-grade-level posttest 2	41.14 (3.44)	41.15 (3.36)	40.33 (3.24)	43.50 (.71)	44.00 (4.24)
Off-grade-level posttest 1	33.00 (4.24)	32.69 (2.10)	32.53 (3.16)	33.00 (1.41)	33.00 (2.83)
Off-grade-level posttest 2	34.29 (3.40)	31.77 (2.62)	31.67 (2.58)	*	35.50 (3.54)
AIMSweb					
Pretest	47.86 (8.30)	44.38 (12.64)	47.27 (11.14)	30.50 (3.54)	47.00 (11.31)
On-grade-level posttest 1	55.29 (9.25)	50.38 (12.93)	52.07 (11.70)	40.50 (14.85)	62.00 (2.83)
On-grade-level posttest 2	58.57 (7.04)	55.38 (10.85)	57.80 (8.14)	41.50 (14.85)	58.50 (9.19)
Off-grade-level posttest 1	29.86 (5.98)	29.15 (10.42)	29.33 (9.42)	22.50 (9.19)	32.50 (2.12)
Off-grade-level posttest 2	38.43 (4.12)	31.77 (6.53)	33.87 (6.24)	27.50 (4.95)	42.00 (5.66)

Note. The score for 1 student of "Other" ethnicity is not reported in this table. Off-grade-level = assessments at the next grade level; * = student scores are the same.

REFERENCES CITED

Achieve Inc. (2010). *On the road to implementation: Aligning assessments with the common core state standards*. Retrieved from http://www.achieve.org/files/CCSS&Assessments.pdf

AIMSweb national norms table: Math computation (2012). Retrieved from https://aimsweb.pearson.com/Report.cfm?inc=rReport18.cfm&ClearSchool=1&Aggregate=1

Alonzo, J., Lai, C., & Tindal, G. (2009). *The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 3* (Technical Report No. 0902). Retrieved from http://www.brtprojects.org/publications/technical-reports

Alonzo, J., Lai, C., & Tindal, G. (2009). *The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 4* (Technical Report No. 0903). Retrieved from http://www.brtprojects.org/publications/technical-reports

Alonzo, J., Lai, C., & Tindal, G. (2009). *The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 5* (Technical Report No. 0901). Retrieved from http://www.brtprojects.org/publications/technical-reports

Alonzo, J., & Tindal, G. (2010). *Teachers' manual for regular easyCBM: Getting the most out of the system*. Retrieved from http://www.brtprojects.org/publications/technical-reports

Babbie, E. (2010). *The practice of social research.* Belmont, CA: Cengage Learning.

Becker, H. J. (1992). 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. doi:10.1016/0883-0355(94)90024-8

Blanton, W. E., Moorman, G. B., Hayes, B. A., & Warner, M. L. (1997). Effects of participation in the fifth dimension on far transfer. *Journal of Educational Computing Research, 16,* 371-398. doi: 10.2190/0YAW-FYAN-2T2B-0LP3

Bunce, B. (2010). *Archdiocese of Portland in Oregon fact sheet 2010-2011*. Retrieved from Archdiocese of Portland website: http://www.archdpdx.org/stats/

Campbell, D., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research*. Boston, MA: Houghton Mifflin.

easyCBM progress monitoring score interpretations guidelines (2011). Retrieved from http://easycbm.com/static/files/pdfs/info/ProgMonScoreInterpretation.pdf

Fuchs, L.S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., … Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology, 98*, 29-43. doi:10.1037/0022-0663.98.1.29

Fuchs, L. S., Fuchs, D., Stuebing, K., Fletcher, J. M., Hamlett, C. L., & Lambert, W. (2008). Problem solving and computational skill: Are they shared or distinct aspects of mathematical cognition? *Journal of Educational Psychology, 100*, 30-47. doi:10.1037/ 0022-0663.100.1.30

Fraenkel, J. R., & Wallen, N. E. (2003). *How to design and evaluate research in education* (5th ed.). Boston, MA: McGraw-Hill.

Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). *Assisting students struggling with mathematics: Response to Intervention (RtI) for elementary and middle schools* (NCEE 2009-4060). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed. gov/ncee/wwc/pdf/practiceguides/ rti_math_pg_042109.pdf

Irish, C. (2002). Using peg- and keyword mnemonics and computer-assisted instruction to enhance basic multiplication performance in elementary students with learning and cognitive disabilities. *Journal of Special Education Technology, 17,* 29-40. Retrieved from EBSCO*host*

Jackson, S. L. (2010). *Statistics plain and simple*. Belmont, CA: Cengage Learning.

Kulik, C.-L. C., & Kulik, J. A. (1991). Effectiveness of computer-based instruction: an updated analysis. *Computers in Human Behavior, 7,* 75-94. Retrieved from EBSCO*host*

Kulik, J. A. (2003). *Effects of using instructional technology in elementary and secondary schools: What controlled evaluation studies say* (SRI Project No. P10446.001). Arlington, VA: SRI International. Retrieved from http://www.sri.com/policy/csted/ reports/sandt/it/

Li, Q., & Ma, X. (2010). A meta-analysis of the effects of computer technology on school students' mathematics learning. *Educational Psychology Review, 22,* 215-243. doi: 10.1007/s10648-010-9125-8

Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research, 71,* 449–521. doi: 10.3102/00346543071003449

Martindale, T., Pearson, C., Curda, L. K., & Pilcher, J. (2005). Effects of an online instructional application on reading and mathematics standardized test scores. *Journal of Research on Technology in Education, 37,* 349-360. Retrieved from http://www.eric.ed.gov/PDFS/EJ690977.pdf

McBurney, D. H., & White, T. L. (2010). *Research methods* (8th ed.). Belmont, CA: Cengage Learning.

McDonald, D., & Schultz, M. M. (2010). *United States catholic elementary and secondary schools 2009-2010*: *The annual statistical report on schools, enrollment and staffing.* Retrieved from: http://www.ncea.org/news/Annual DataReport.asp

Means, B. (1994). *The technology and education reform: The reality behind the promise*. San Francisco, CA: Jossey-Bass.

Means, B. (2008). Technology's role in curriculum and instruction. *The SAGE handbook of curriculum and instruction*. Retrieved from http://sage-ereference.com/hdbk_ curriculum/Article_n7.html

National Assessment of Educational Progress (2009). *The nation's report card. Classroom context: Computer use*. Retrieved from http://www.nationsreportcard. gov/math_2009/context_4.asp

National Assessment of Educational Progress (2010). *The nation's report card. What does the NAEP mathematics assessment measure?* Retrieved from http://nces.ed.gov/nationsreportcard/mathematics/whatmeasure.asp

National Council of Teachers of Mathematics (2009). *Executive summary: Principles and standards for school mathematics*. Retrieved from http://www.nctm.org/ uploadedFiles/Math_Standards/12752_exec_pssm.pdf

National Mathematics Advisory Panel (2008). *Foundations for success: The final report of the national mathematics advisory panel*. Retrieved from http://www2.ed.gov/ about/ bdscomm/list/mathpanel/report/final-report.pdf

Nese, J. F. T., Lai, C.-F., Anderson, D., Jamgochian, E. M., Kamata, A., Saez, L., …Tindal, G. (2010). *Technical adequacy of the easyCBM mathematics measures: Grades 3-8, 2009-2010 Version* (Technical Report No. 1007). Retrieved from http://www.brtprojects.org/ publications/technical-reports

Overton, T. (2000). *Assessment in special education: An applied approach.* Upper Saddle River, NJ: Prentice-Hall.

Pearson Education. (2008) AIMSweb [Web-based data management system]. Retrieved from http://www.aimsweb.com

Pearson Education. (2010). *AIMSweb Mathematics Computation Administration and Technical Manual*. Bloomington, MN: *Author*.

Quinn, D. W., & Quinn, N. W. (2001). *Evaluation series: Grades 1-8, Apache Junction Unified School District 43, Apache Junction, Arizona.* Retrieved from PLATO Learning, Inc. website: http://www.plato.com/media/Evaluation%20Studies/A/Apache%20Junction%20Unified%20School%20District.pdf

Renaissance Learning, Inc. (2009). *MathFacts in a Flash*: *When you build a strong math foundation, students really soar*. Wisconsin Rapids, WI: *Author*.

Royer, J. M., Greene, B. A., & Anzalone, S. J. (1994). Can U.S. developed CAI work effectively in a developing country? *Journal of Educational Computing Research, 10,* 41-61. doi: 10.2190/PMHQ-PYCD-2JB7-PLB8

Salerno, C. A. (1995). The effect of time on computer-assisted instruction for at-risk students. *Journal of Research on Computing in Education, 28,* 85-97. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ516557&login.asp&site=ehost-live&scope=site

Skinner, B.F. (1989). Teaching machines [Supplemental material]. *Science, 243*, 1,535. doi:10.1126/science.243.4898.1535-b

Slavin, R. E., & Lake, C. (2008). Effective programs in elementary mathematics: A best-evidence synthesis. *Review of Educational Research, 78,* 427-515. doi: 10.3102/0034654308317473

Stockburger (1996). Analysis of crossed designs. In *Introductory statistics: Concepts, models, and applications WWW version 1.0*. Retrieved from http://www.psychstat.missouristate. edu/introbook/sbk00.htm

Tomlinson, C. (1999). *The differentiated classroom: Responding to the needs of all learners*. Alexandria, VA: Association for Supervision and Curriculum Development.

Thurber, R. S., Shinn, M. R., & Smolkowski, K. (2002). What is measured in mathematics tests? Construct validity of curriculum-based math measures. *School Psychology Review, 31,* 498-513. Retrieved from http://www.aimsweb.com/uploads/news/id30/thurber_et_al.pdf

Trochim, W. M. K. (2006). *External validity*. Retrieved from http://www.socialresearchmethods.net/kb/external.php

Wittman, T. K., Marcinkiewicz, H. R., & Hamodey-Douglas, S. (1998). *Computer assisted automatization of multiplication facts reduces mathematics anxiety in elementary school children* (Report No.143). Retrieved from ERIC Ebsco*host*

Xin, J. F. (1999). *Computer-assisted cooperative learning in integrated classrooms for students with and without disabilities*. Retrieved from http://find.galegroup.com/gtx/infomark.do?&contentSet=IAC-Documents&type=retrieve&tabID=T002&prodId=CDB&docId=A72984160&source=gale&srcprod=CDB&userGroupName=s8492775&version=1.0

Ysseldyke, J. E., Thill, T., Pohl, J., & Bolt, D. (2005). Using *MathFacts in a Flash* to enhance computational fluency. *Journal of Evidence Based Practices for Schools,6*, 59-89.