



Michigan Technological University
Create the Future Digital Commons @ Michigan Tech

Dissertations, Master's Theses and Master's
Reports - Open

Dissertations, Master's Theses and Master's
Reports

2011

Attempting to improve standardized test results using Study Islands' web-based mastery program

Paul J. Dube'
Michigan Technological University

Follow this and additional works at: <https://digitalcommons.mtu.edu/etds>



Part of the [Science and Mathematics Education Commons](#)

Copyright 2011 Paul J. Dube'

Recommended Citation

Dube', Paul J., "Attempting to improve standardized test results using Study Islands' web-based mastery program ", Master's report, Michigan Technological University, 2011.
<https://doi.org/10.37099/mtu.dc.etds/524>

Follow this and additional works at: <https://digitalcommons.mtu.edu/etds>



Part of the [Science and Mathematics Education Commons](#)

**ATTEMPTING TO IMPROVE STANDARDIZED TEST RESULTS
USING STUDY ISLANDS' WEB-BASED MASTERY PROGRAM**

By

Paul J. Dube'

A REPORT

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

(Applied Science Education)

MICHIGAN TECHNOLOGICAL UNIVERSITY

2011

©2011 Paul J. Dube'

This report, "Attempting To Improve Standardized Test Results Using Study Islands' Web-Based Mastery Program," is hereby approved in partial fulfillment of the requirements for the Degree of MASTER OF SCIENCE IN APPLIED SCIENCE EDUCATION.

Department of Cognitive and Learning Sciences

Signatures:

Advisor

Shari Stockero

Committee Member

Susan Amato-Henderson

Committee Member

Yolanda Munoz Maldonado

Department Chair

Brad Baltensperger

Date

Abstract

State standardized testing has always been a tool to measure a school's performance and to help evaluate school curriculum. However, with the school of choice legislation in 1992, the MEAP test became a measuring stick to grade schools by and a major tool in attracting school of choice students. Now, declining enrollment and a state budget struggling to stay out of the red have made school of choice students more important than ever before. MEAP scores have become the deciding factor in some cases.

For the past five years, the Hancock Middle School staff has been working hard to improve their students' MEAP scores in accordance with President Bush's "No Child Left Behind" legislation. In 2005, the school was awarded a grant that enabled staff to work for two years on writing and working towards school goals that were based on the improvement of MEAP scores in writing and math. As part of this effort, the school purchased an internet-based program geared at giving students practice on state content standards.

This study examined the results of efforts by Hancock Middle School to help improve student scores in mathematics on the MEAP test through the use of an online program called "Study Island." In the past, the program was used to remediate students, and as a review with an incentive at the end of the year for students completing a certain number of objectives. It had also been used as a review before upcoming MEAP testing in the fall. All of these methods may have helped a few students perform at an increased level on their standardized test, but the question remained of whether a sustained use of the program in a classroom setting would increase an understanding of concepts and

performance on the MEAP for the masses. This study addressed this question.

Student MEAP scores and Study Island data from experimental and comparison groups of students were compared to understand how a sustained use of Study Island in the classroom would impact student test scores on the MEAP. In addition, these data were analyzed to determine whether Study Island results provide a good indicator of students' MEAP performance. The results of the study suggest that there were limited benefits related to sustained use of Study Island and gave some indications about the effectiveness of the mathematics curriculum at Hancock Middle School. These results and implications for instruction are discussed.

Acknowledgements

Dedicated to my wife Connie, daughter Payton, and son Dylan for being supportive and understanding of my time and effort.

It is also dedicated to my 7th grade Math students who took part in this study during the 2009-2010 school year. Without their efforts this project would not have been possible.

Table of Contents

Abstract	iii
Acknowledgements	v
List of Tables	ix
List of Figures	x
Chapter 1: Introduction	1
Adoption of Study Island	1
Uses of Study Island	3
The Study	4
Chapter 2: Literature Review	7
Legislation Mandating State Standards and Assessments	7
Alignment	8
Technology in the Classroom	9
Using the Internet to Supplement Mathematics Instruction	14
Virtual Parent Involvement	16
Summary	17
Chapter 3: Methodology	19
Participants	19
Study Island Program	20
General Structure	20
Student Use of Study Island	21
Game Mode	22
Teacher Reports	23
Procedure	24

Pre-MEAP Activities	24
Use of Study Island with the Experimental Group	25
Use of Study Island with the Comparison Group	31
Post-test.....	32
Data Collection and Analysis.....	32
MEAP Results.....	32
Study Island Records	34
Researcher Journal	36
Chapter 4: Results	37
Data Summary	37
Comparison of Groups at Start of Study.....	39
MEAP Test Results.....	39
Changes In MEAP Scores.....	39
MEAP Scores By Level	40
Grade Level Content Standard Breakdown	46
Study Island Data.....	48
MEAP / Study Island Correlation	50
Study Island Post-test and 2010 MEAP Score Analysis.....	50
Time Spent on Study Island and 2010 MEAP Score Analysis	54
Study Island Blue Ribbon Percent and 2010 MEAP Score Analysis	54
Assessment of Curriculum.....	55
Chapter 5: Conclusion and Implications.....	57
Limitations of Study	61
Suggestions	63

References.....	64
Appendix A – IRB Documents	67
Appendix B – Data Tables.....	70

List of Tables

Table 1: Participant Summary Data for MEAP and Study Island Tests.....	37
Table 2: Experimental Group Summary Data for MEAP and Study Island Tests	38
Table 3: Comparison Group Summary Data for MEAP and Study Island Tests	38
Table 4: Hancock Middle School 2010 Mathematics MEAP Results By Performance Level	41
Table 5: Hancock Middle School 2009 Mathematics MEAP Results By Performance Level	41
Table 6: Hancock Middle School 2010 Mathematics MEAP Level Movement for Experimental Group	43
Table 7: Hancock Middle School 2010 Mathematics MEAP Movement for Comparison Group	43
Table 8: Hancock Middle School 2010 Mathematics MEAP Result Changes (Including Changes in Level Subcategories)	44
Table 9: Hancock Middle School 2010 Mathematics MEAP Results Performance of Level 1 Students (Including Changes in Level Subcategories)	45
Table 10: Hancock Middle School 2010 Mathematics MEAP Results Performance of Level 2 Students (Including Changes in Level Subcategories)	45
Table 11: Hancock Middle School 2010 Mathematics MEAP Results Performance of Level 3 Students (Including Changes In Subcategories)	46
Table 12: Hancock Middle School 2010 Mathematics MEAP Results By Grade Level Content Standard.....	48
Table 13: Hancock Middle School Study Island and MEAP Results By Grade Level Content Standard.....	53

List of Figures

Figure 1: Study Island and MEAP Question Comparison	10
Figure 2: State of Michigan GLCE Breakdown	20
Figure 3: Relationship between 2010 MEAP Scores and Study Island Post-test Scores ..	52
Figure 4: Relationship between Experimental Group 2010 MEAP Scores and Study Island Blue Ribbon Performance	55

Chapter 1: Introduction

Standardized testing in the state of Michigan has always been a tool to measure a school's performance and to help evaluate school curriculum. However, with the school of choice legislation in 1996 (State of Michigan, 2011), the Michigan Education Assessment Program (MEAP) test also became a means by which to grade schools and, thus, a school's MEAP scores became a major tool in attracting school of choice students. Now, 15 years later, declining school enrollments and a state budget that is struggling to stay out of the red have made school of choice students more important to school districts than ever before. Many schools have turned to advertising in local newspapers to lure prospective students, and MEAP scores have become the deciding factor in some cases.

Adoption of Study Island

For the past five years Hancock Middle School has been working hard to improve their MEAP scores in accordance with President Bush's No Child Left Behind Act of 2001 (U.S. Congress, 2001). In September 2005, the school was awarded a Comprehensive School Reform grant that enabled staff to work for two years in guided professional development focused on writing and working towards school goals that were based on the improvement of MEAP scores in writing and mathematics. School staff and administrators worked with consultants and mentors to develop activities that would help integrate cross-curricular writing activities into teachers' instruction and to consider ways to increase awareness of the importance of the MEAP testing and change the attitudes about this testing of faculty and students alike. Along with changing attitudes about the testing, the school purchased a web-based program that is a standards-based program

geared at giving students practice answering MEAP-like questions related to state content standards.

The web-based mastery program used in this study is called Study Island. Study Island (2009) is the flagship of Archipelago Learning, Inc., which also produces other materials for K-12 schools. Study Island was first developed for the Ohio Proficiency Test Program, and Ohio and Illinois were the first states to have web-based subscriptions to the program. The company has seen the success of their program grow in the last ten years, as it is now used in all 50 states and three provinces in Canada. As of the 2009-2010 school year, the company reported that in the U.S., nearly 22,000 schools and approximately 10 million students used Study Island. In Michigan, as of September 2010, there were 258 school districts, 585 schools, 443,000 students, 33,000 teachers, and 250 individual subscribers to Study Island (personal communication with Archipelago Learning employee, September 15, 2010).

Study Island seemed to be a natural fit for the Comprehensive School Reform Grant. The administration at Hancock Middle School received electronic promotional ads from the company that piqued their interest, as the capabilities of the Study Island program seemed to fall in line with the goals of the Comprehensive School Reform Grant. Study Island offered a free trial of their services, which allowed administrators to see how it might work in the district. Further consideration was given to purchasing the web-based program after administrators had done some research by contacting other schools in the area and around the state that were using the program. The research showed that schools were using the program in different ways and that it seemed to be

flexible to the needs of the district. The grant paid for the first year of service at a cost of about \$1200 dollars for use by students in grades 6-8. The school has continued to use the service of Study Island with generous yearly donations from the district's foundation.

Hancock Middle School principal and now superintendent, Monica Healy, was responsible for bringing Hancock Middle School to the "Island." It was her work that earned the district the Comprehensive School Reform Grant and her guidance that led to the development of three strong goals that the middle school could work to obtain. Those goals include improving mathematics and reading MEAP scores at HMS. In an interview, Ms. Healy (personal communication, May 27, 2010) stated that she had two main goals for Study Island when she first decided to purchase the product. First, when teachers are not available, she wanted to give students a means to review and practice for the MEAP with an easy and accessible tool. Her second, and perhaps more important, goal was to increase students' mathematics and reading scores on the MEAP.

Uses of Study Island

Hancock Middle School uses the Study Island program in a variety of ways to achieve the goal of improving MEAP scores. Communication with parents is the first phase of the plan. At the beginning of each school year, all parents of middle school children are sent a letter informing them about the Study Island program and its benefits to their children. Teachers are encouraged to use Study Island in their classrooms throughout the year. Tutors work with students using Study Island as a tool. In the fall of the school year, before the MEAP test is given, the program is used as a review tool; in the winter, after MEAP scores have been returned to the district, Study Island is used as a

tool for remediation and all teachers give a homework score to students for completing the remedial work. Finally, Study Island has become the curriculum for the school's summer school program. Students are given a series of homework assignments to complete and are then given assessments based on the concepts where deficiencies were noted.

Teachers at Hancock Middle School are encouraged, but not required, to use the web-based mastery program in the classroom. It is up to each individual teacher to decide how much or how little Study Island will be included as part of their curriculum. The only time during the year that teachers are required to use the study program is during the MEAP remediation period during the winter term. Administrators do not monitor teachers' use of the program, but if they choose to do so, they would be able to determine which teachers were regularly using the program by looking at student statistics by grade level and subject, and comparing student use to the instructors that they have for particular subjects.

The Study

Has Study Island been effective in helping Hancock Middle School obtain its goals? Ms. Healy (personal communication, May 27, 2010) thinks so. In her opinion, Study Island is a great program, but more can be done. For example, the program could be used more effectively if its use were more widespread among teachers. A program like Study Island will only be effective if the staff is committed to learning and using the program. If students' use of the program is not supported in the classroom then the likelihood of effective student use will be low, as there will not be any checks and

balances for time spent.

The research study will focus on understanding the effects of the use of Study Island as a way of improving students' scores on the mathematics portion of the MEAP test. In the past, the program was used by teachers to remediate students on topics taught during the previous year and as a review tool for topics taught during the current year, with incentives for students who have completed a certain number of objectives. It has also been used as a "change of pace" by some teachers who periodically want to get students out of the classroom, as well as a tool for review before upcoming MEAP testing in the fall. Students could also access the program from home at their leisure.

I, personally, have used Study Island as a tool to supplement my textbook. With the change in state content standards in 2006 (State of Michigan, 2006) my textbook became less effective, as I found there were new required concepts that it did not adequately address. The Study Island program helped to ensure that I did cover necessary content gaps. I also used the program to generate a year-end final examination to help evaluate how successful I had been with regard to meeting the Michigan Grade Level Content Expectations (GLCEs). The total time spent using the Study Island website would be estimated at less than 10 hours per classroom and 3 hours per student.

While all of the methods used by myself and other teachers may have helped a few students perform at an increased level on their standardized test, it is not known whether a sustained use of the program in a classroom setting could increase students' understanding of concepts, and thus, performance on the mathematics portion of the MEAP, for the masses. In an attempt to understand the potential effects of the Study

Island program, this study focuses on answering the question:

- To what extent did continuous teacher-supported use of the Study Island software in a classroom setting improve student test results on state standardized mathematics testing?

The study also focused on answering the following underlying questions:

- To what extent did students' Study Island post-test scores correlate to their MEAP test results? Was a Study Island post-test an accurate indicator of students' performance on the MEAP test?
- What did the Study Island data indicate about the effectiveness of Hancock Middle School's mathematics curriculum?

Chapter 2: Literature Review

Legislation Mandating State Standards and Assessments

In 1994, the Improving America's Schools Act [IASA] (U.S. Congress, 1994) was passed into law. It was designed to improve student achievement in the U.S. by calling for higher standards for all children, including Title 1 students (Wixson, Fisk, Dutro, & McDaniel, 2001). The IASA required states to develop state-specific standards, along with high-quality carefully aligned assessments that could be used to evaluate student performance (Payzant & Levin, 1995). In Michigan, the assessment tool used to measure students' achievement of the state's content standards—the Grade Level Content Expectations (GLCE's)—is known as the Michigan Education Assessment Program or MEAP. This assessment was developed through a sequential development approach in which the state developed a sequence of standards, frameworks, and assessments that would ensure that the grade level standards were aligned from year to year and were met by all students (Wixson, Fisk, Dutro, & McDaniel, 2001).

More recently, the No Child Left Behind Act (U.S. Congress, 2001) required that all schools nationally have academic standards in the core areas of reading, mathematics, and science. It mandated that states set benchmarks for proficiency standards in these three core areas, with the goal of achieving 100% student proficiency by 2014. The implied idea of NCLB is to reach every student and help them meet the state standards, regardless of any sub-grouping or categorizing, and to personalize instruction for students to enable their success.

Alignment

With schools across the state of Michigan and the country facing tougher legislation related to academic performance, how are schools addressing the national standards for improvement? In Michigan, schools are focusing on enacting a strong curriculum that is aligned in accordance to state and federal standards. What is an aligned curriculum? According to the IASA, “Title 1 made the link between standards and assessment apparent by requiring states to develop assessments that are aligned with the state’s challenging content and performance standards and provide coherent information about student attainment of such standards” (U.S. Congress, 1994, p. 8). According to the Study Island Scientific Research Base, the term alignment can have many different meanings, but they follow a definition from Webb (1997), who defines alignment within educational settings as “the degree to which expectations and assessments are in agreement and serve in conjunction with one another to guide the system toward students learning what they are expect to know and do” (Magnolia Consulting, 2008, p. 2).

The State of Michigan looks at alignment in terms of test validity. In the state’s technical report, validity refers to “the extent to which a test measures what it is intended to measure and how well it does so” (Michigan Department of Education, 2008, p. 143). Since 1970, the State of Michigan has put a lot of money and effort into documenting student achievement and the success of school districts by analyzing the MEAP tests for validity annually as changes to the assessment have occurred in response to changes in the content standards.

Study Island is a web-based mastery program designed to help schools with state specific content and is aligned with each state's curriculum standards. Study Island program authors developed the content of Study Island from an in-depth analysis of each state's learning objectives to create highly specific and individualized versions of the program for each state (Magnolia Consulting, 2008). To illustrate the alignment of the Michigan version of Study Island with Michigan's GLCEs and to give the reader a snapshot of the types of questions that students are facing on the MEAP, Figure 1 compares a number of released items from the MEAP (State of Michigan, 2008; 2007; 2006) to problems from the Study Island website (www.studyisland.com). In addition, the corresponding Michigan 7th grade GLCEs that the questions are meant to address are included. This is just a sample of some of the items that students are assessed on and not a complete comparison. However, it allows the reader to see the strong resemblances from one assessment to the next. Being able to look at the relationships between questions on the state assessment and the web-based mastery product helps to judge just how close the questioning and thinking that goes into solving each question compares between the two assessments.

Technology in the Classroom

The use of computers in school classrooms grew drastically in the 1980's. Twenty-five percent of all schools in the nation in the early 80's used computers in their instruction, as compared to almost 100% of all schools by 1990 (Cotton, 2007). Computer use, however, does not necessarily mean improved learning. The most

important question, then, is has the growth and use of computer technology helped to increase the learning potential of students or is it just a fascinating form of entertainment?

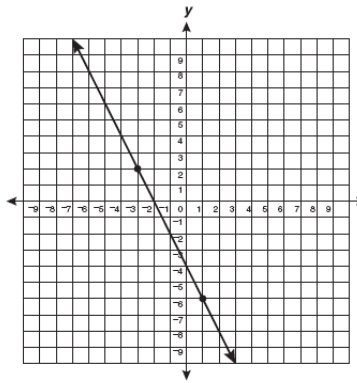
Michigan GLCE	SI Example	MEAP Example												
N.MR.07.02 Solve problems involving derived quantities such as density, velocity, and weighted averages.	Mario rode 200 miles on a passenger train. It took him 5 hours to reach his destination. What was the average speed of the train?	The population of Michigan was 10,079,985 in 2003. The area of Michigan is 59,685 square miles. What was the approximate population density of people per square mile in 2003?												
A.PA.07.06 Calculate the slope from the graph of a linear function as the ratio of “rise/run” for a pair of points on the graph, and express the answer as a fraction and a decimal; understand that linear functions have slope that is a constant rate of change.	Determine the slope of the line that passes through the points (1, -1) and (4, -3).	Which appears to be the slope of the line graphed on the grid below? 												
D.AN.07.03 Calculate and interpret relative frequencies and cumulative frequencies for given data sets.	The weights of 100 7th graders are shown in the table below: <table border="1" data-bbox="649 1253 946 1568"><thead><tr><th>Weights in Pounds</th><th>Number of Students</th></tr></thead><tbody><tr><td>71 - 80</td><td>7</td></tr><tr><td>81 - 90</td><td>13</td></tr><tr><td>91 - 100</td><td>32</td></tr><tr><td>101 - 110</td><td>36</td></tr><tr><td>111 - 120</td><td>12</td></tr></tbody></table> What is the cumulative frequency of students weighing less than 101 pounds?	Weights in Pounds	Number of Students	71 - 80	7	81 - 90	13	91 - 100	32	101 - 110	36	111 - 120	12	A survey was given to 250 women. Each woman was asked how many children she had. The results showed that 0.28 of the women had 3 or more children. How many of the women surveyed had fewer than 3 children?
Weights in Pounds	Number of Students													
71 - 80	7													
81 - 90	13													
91 - 100	32													
101 - 110	36													
111 - 120	12													

Figure 1. Study Island and MEAP Question Comparison

According to Cotton (2007), research has demonstrated that Computer-Assisted Instruction (CAI) has benefitted education and student achievement alike. She states that research has found that CAI helps all students improve their learning, but seems to have a larger effect on younger (elementary) as compared to older (high school/college) students. Cotton's research indicates that supplementing a typical classroom lecture with the use of CAI can produce better learning results. She goes on to say that CAI has been shown to help increase learning rates, retention, and improve student attitudes and that students who were introduced to CAI demonstrated an increased learning potential when compared to those who were not. Particularly relevant for schools trying to help their students achieve the NCLB legislation goals, Cotton points out that research has proven that CAI can help lower achieving students gain success. Many lower achieving students lack self-confidence and avoid being singled out. CAI programs can assist these students by having such things as privacy, immediate feedback, and repetition built into the instructional activities.

In a study specific to mathematics, Ash (2004) examined whether CAI helped increase middle school students' math scores. His study consisted of using CAI for 1 hour a week throughout the school year. In this study it was found that the one-hour a week of CAI was beneficial, as the students who used the CAI scored better on a post-test than those who only had traditional instruction.

Since the introduction of computers and television as early forms of educational technology, technological applications in the classroom have taken off and have become a major component of education today. D'Arcy, Eastburn, and Bruce (2009) set out to see

how different types of media can address diverse student needs. In their study, they classified students into four distinct types of learners based on the Gregorc Style Delineator (Gregorc, 1984). According to the Gregorc Style Delineator the four styles are based on perceptual ability and ordering ability. Perceptual ability is broken down into two groups, abstract and concrete, and ordering abilities are defined by random and sequential ability. By combining the perceptual and ordering ability, the four groups created consist of abstract random, concrete random, abstract sequential, and concrete sequential abilities. D'Arcy, Eastburn, and Bruce's study consisted of presenting multiple instructional media types to students in collegiate classes to see if students preferred or learned better from certain types of instructional media. The study concluded that there were eight types of instructional media that were rated effective by students, regardless of learning styles. These included the use of overheads, PowerPoint presentations, i-clickers®, and on-line quizzes.

In the study, effective use of overheads—digitized and displayed via LCD projector when diagrams and drawings were being discussed—were noted by students as positively affecting learning (D'Arcy, Eastburn, & Bruce, 2009). PowerPoint slides were also noted as effective for presenting instructional images. Students responded well to these forms of media because they helped them visualize what the instructor was talking about and, thus, aided in their understanding; they also helped the instructor organize material in an easy-to-follow format.

The i-clicker generated the greatest positive response from students. An i-clicker is a remote control that students use to interact with a computer. The clicker provides

instructors the opportunity to create a game-type scenario, where students or groups of students are given a question with multiple-choice answers. The students select their answers using the remote control and the answers are tabulated on a computer and displayed on a projection screen. Because the i-clicker is an active learning medium, students found it helpful because it gave them a chance to test their knowledge before, during, and after particular topic discussions in class (D'Arcy, Eastburn, & Bruce, 2009). Students also liked that the use of the i-clicker was private and indicated that it was fun. The i-clicker also provided immediate feedback to the teacher.

Internet-based quizzes were reported to help students become more responsible and active in their learning. By taking the online quizzes, students felt that it forced them to stay current with classroom material, gave students an idea of what was most important from a section of instruction, and gave immediate feedback so students tended to go back over what they did not understand (D'Arcy, Eastburn, & Bruce, 2009).

Because Study Island (SI) is a web-based program, it can be used as an alternative media like the ones seen in D'Arcy, Eastburn, & Bruce's (2009) study. Additionally, SI contains many of the very media components that the study found to be effective. In particular, the program includes PowerPoint presentations for some lessons, notes for all sections that students can access online or teachers can put display using an LCD projector, online quizzes that give instant feedback and the ability for instructors to use i-clickers in their classrooms for the multiple choice format questions.

In general, D'Arcy, Eastburn, & Bruce's (2009) study lends prudence to the value of diverse media and the benefits to the students who are introduced to them, regardless

of the type or style of learning they may possess. For these media to be successful, however, the instructors must be trained and versed in using them and willing to become familiar with alternative teaching strategies. To help in this regard, Study Island also offers online webinars to help train instructors on its uses.

Using the Internet to Supplement Mathematics Instruction

In the electronic world of today, teaching is sometimes as much about entertaining as it is about instructing, so in order to help students become proficient in mathematics it is important to access and use as many different tools as possible (Kilpatrick, Swafford, & Findell, 2001). One of those resources is the internet. The internet can be a positive educational tool, as teachers are able to find and use materials that help students learn the content. For example, there are many sites on the world-wide-web that can help teachers develop lessons that address learning objectives in a more engaging way. Things like online manipulatives, video from sites such as YouTube (www.youtube.com), explanations and notes for topics, PowerPoint presentations, practice problems, and online calculators can be used for classroom purposes, often free of charge.

Study Island is one of these online educational tools that can be used by teachers to improve student learning. Although it is not free, the SI developers claim that the cost is worth the results. For example, the 2007 report for Michigan—which included data for schools that started using the Study Island program during the school year of 2005-2006—noted that the percentage of students meeting or exceeding state math standards in

schools using Study Island increased 13.08%, from 61.89% to 72.70%, while scores statewide increased by only 9.90% (Study Island, 2008). Furthermore, they report,

In 2006, 96% of Michigan *Study Island* schools met Adequate Yearly Progress while only 85% of other Michigan schools met AYP. Study Island schools met AYP at an 11% higher rate than non-Study Island schools in Michigan. Additionally, when compared to other Michigan schools, *Study Island* schools had a 38% higher rate of improvement for students that did not meet standards on the MEAP math assessment (Magnolia Consulting, 2008).

Although the use of internet-based tools such as Study Island can benefit classroom instruction, there are also some drawbacks associated with the use of the internet in a classroom. Some critics of the internet say that the teacher becomes obsolete and the students are the ones who are producing all the work (Loong & White, 2004). Another drawback identified by Loong & White (2004) in a survey of teachers was the issue of the time it takes to plan a lesson involving technology and the time it typically takes to integrate the computer into the classroom. When students are allowed the opportunity to interact with the technology, time needs to be added to the lesson to both allow students to investigate ideas and to communicate what they have discovered from their inquiries. Other notable issues come in the form of reliability of the technology and having the support necessary when technological issues arise, as well as the availability of the computer room or computers in general. Finally, any resource on the world-wide-web that offers solutions or answers to mathematics problems poses a challenge to teachers, as students may come back with answers but not the knowledge to support the work that they have.

Interactions between the teacher, the students, and the curriculum are the basis for supporting student learning, and how these components come together significantly affects the development of students' understanding of the topics (Kilpatrick, Swafford, & Findell, 2001). Teachers may use the internet to allow students to investigate a topic but without allowing those students the opportunity to discuss what they were working on or their thinking about the activity, the teacher may not ever fully know the level of students' comprehension (Kilpatrick, Swafford, & Findell, 2001). Therefore, if a teacher is not trained to facilitate productive discussions, to recognize important ideas in student thinking, and assess their understanding while discussing the math behind an activity, then using the internet as a tool for instruction may not be very effective in supporting student learning. When used well, however, internet-based instruction can be implemented in a way that effectively blends the three components of teacher, students, and curriculum.

Virtual Parent Involvement

Magnolia Consulting (2008) reports that "67% of American households with school-aged children have a computer connected to the internet" (p. 15) and via the internet, schools are making more information available to parents. In fact, in a survey of parents with children in grades K-8, parents indicated a greater use of the internet to stay connected with their child's school than other means of correspondence, such as phones. From parental email to school websites, school districts are trying to reach out to parents and make it easier for them to stay informed. Many schools offer websites with daily announcements, daily homework and other pertinent school information. Teachers, as

well, are progressing to the technological age by offering their own pages with wikis, notes, videos, and more to help students and parents stay caught up with their class work. Some schools are also using grading software such as Skyward (Skyward, Inc., 2011) that allows parents and students alike the opportunity to check on daily progress from homework scores to overall grades.

Parents' communication with the school during their children's adolescent years was found to produce positive results. Bouffard (2006) points out that children whose parents regularly communicated with school personnel tended to have better grades, better test scores, and greater graduation rates. According to Bouffard, the use of the internet to communicate with school is positively related to educational expectations, meaning that more correspondence between schools and parents usually indicates a great academic expectation by these parents. According to SI research "If parents have higher expectations for their children, they may be more apt to monitor their children's ongoing progress" (Magnolia Consulting, 2008, p.16). SI allows parents to monitor student work online, which makes them more aware of their child's math skills. This awareness may invite parents to communicate with the teacher to address deficiencies early in the school year, rather than waiting until the student has had substantial problems in their math class, as is often the case.

Summary

Based on the literature reviewed, two main concepts evolved. The first was the idea that CAI is beneficial. The use of CAI can be a positive component of classroom instruction, provided the instructor is knowledgeable about the technology. It is important

that the use of the internet is not used as an introduction to computers or as free day for students, but rather, that it is used to support the instructional goals. Second, CAI is only as good as the curriculum it is used to enhance. When using CAI as a tool for instruction, it should be tied to meaningful content taught in class. Therefore, it is important that instruction and use of technological tools are tied to an aligned curriculum. It takes time to develop lessons and integrate the use of technology, as well as time to operate the technology and discuss outcomes, so it is important that the time is well spent. If the curriculum is lacking, the CAI may have some positive effects for the students but the end may not outweigh the means.

Study Island is a web-based mastery program that is used by districts and schools to help improve state standardized testing. Study Island makes claims of their success and has statistical data to back it up. This study looks to challenge or support these claims while using Study Island as a regular CAI tool to teach 7th grade mathematics, and to determine whether the use of Study Island is worth the time, money, and effort for schools, teachers, and students.

Chapter 3: Methodology

Participants

Participants in the study were students enrolled in one of three seventh grade mathematics classes taught by the researcher at Hancock Middle School during the 2009-2010 academic year. One class was assigned to be the experimental group, and the other two classes became the comparison group. Students were assigned to one of the mathematics classes by a school administrator, so it is assumed that the groups were comparable. This assumption will be verified using pre-test results, however.

All of the students were informed of their rights as human subjects and both they and their parents were given the opportunity to sign an informed consent form prior to participating in the study (MTU IRB protocol M0490; see Appendix A for IRB approval form and participant consent letters). Two students chose not to participate, ten consent forms were not returned, and a couple of students moved out of state during the study period, resulting in a pool of 49 participants. The first hour class, with 14 students participating in the study, was chosen to be the experimental group that would receive extra instruction using Study Island. This class was chosen out of necessity, since first hour was the only hour of the day during which consistent access to a computer lab was guaranteed. Computer and internet access were vital tools in this research, since students needed to use them for accessing the web-based Study Island program. This left the two other class sections as the control group, with 35 participating students.

Study Island Program

General Structure

Study Island is developed to integrate state curriculum standards. Since every state develops their own curriculum, Study Island has developed a way to address each states needs. Since Study Island is state specific, it focuses on Michigan's four main Grade Level Content Expectations (GLCEs) for middle school students by grade level: Number and Operations, Algebra, Geometry, and Data and Probability. Each of these four main content areas is broken down into *tasks*, which are defined as a set of questions related to a specific topic within a content area. For example, tasks in Number and Operation are related to topics that include derived quantities such as velocity and density, as well as working with rates, ratios, proportions, and finding square and cube roots. Within Study Island, there are 18 total tasks that are related to the four categories mentioned above (see Figure 2).

State of Michigan 7th Grade Level Content Expectations (GLCE's)			
Number and Operations	Algebra	Geometry	Data and Probability
Derived Quantities	Linear Relationships	Similar Figures	Interpret Graphs
Rates, Ratios, and Proportions	Linear Functions	Similar Triangles	Scatter Plots
Square & Cube Roots	Inversely Proportional Relationships		Frequencies
Compute Solutions	Properties of Arithmetic		Central Tendency
Estimate Solutions	Algebraic Expression Arithmetic		
Real World Problems	Solve Linear Equations		

Figure 2. State of Michigan GLCE Breakdown

Student Use of Study Island

To work on the program, students log in to Study Island, select a topic or topics from the list of 18, and select the number of problems they wish to work on. They are then given random computer generated problems from the topic or topics in multiple-choice form. One problem at a time appears on the screen. The student selects an answer from four, and occasionally five, potential solutions. If the student is wrong, he or she can continue to select an answer until it is right, but the problem will be scored as incorrect. The student is then given a choice whether to view an explanation of the problem. A student can quit the activity at any time and only the questions answered will be credited to their score.

Students are awarded blue ribbons for performing tasks with a minimum level of accuracy. For each task that is passed, a blue ribbon is awarded to students who score between 60% and 70%. This percent varies by the task. Students have to complete a minimum of 10 to 15 problems related to a specific topic, which also varies by task, to be awarded a ribbon. The program is adaptive to student performance, with more proficient students needing to complete fewer problems to pass a task. For example, a student who is more proficient in a particular area may earn a ribbon in 10 or 15 questions, and another student may take 35 to 50 or more questions to earn the same ribbon. This means that the students who need more practice in a particular area are required to complete more problems. The program keeps a record of problems attempted and those answered correctly. As a result, the student's score for a specific task is continuously changing and

a student can work towards a blue ribbon over time if they do not complete all of the problems in a single session.

When a student is struggling with a specific problem, the program will give a building block activity that must be completed by the student before moving on. This activity is an added topic designed to help improve skills needed to successfully complete the original problem. Once the building block topic is passed, a student can go back and complete the original problem with which they struggled. For example, if a student is struggling with a problem about solving a linear equation, he or she may be given a building block topic on Solving Linear Equations. A sample building block question might look similar to the following: *Brittany gives piano lessons at the Music Conservatory. A 2.5 hour lesson costs \$75.00. What is the price per hour?* If the student successfully completes the building block, a white ribbon will be awarded. If the student does not pass the building block, he or she receives a red triangle that tells them to move on to the next topic, and more help is needed with the topic.

Game Mode

Study Island also includes a game mode, with access granted by the school administrator. Students are able to access one of several games that present problems to students from the topics they choose or, if a student desires, random topics. The games are presented with visual graphics and questions are posed in a game format such as baseball. In baseball, if the student gets a question right, a base is earned. An incorrect answer is an out. The student tries to score as many runs as he or she can. Students are able to compete against all Study Island users across their state and if they score a high

mark, it is added to the Hi-score page for all to see, just like an arcade game. For this study, the game mode was not used in class, but students did have access to play the games on their own.

Teacher Reports

The software provides teachers and administrators many opportunities to check student activity. The first type of report that a teacher may use is called a class grade book report. This report allows the teacher to look at student activity by number of login sessions, time spent online, number of questions answered and the number correct, percentage correct, and MEAP grade. If a student's name is selected, the report will link to an individual student report that includes a breakdown of the student's activities and level of success on all 18 topics. From this point, a tab allows the teacher to compare the current student to other students. This feature allows an instructor to identify students who may be struggling and whether there are common topics that are difficult for many students. This would allow a teacher to make an individualized or whole class plan for instruction or remediation.

Another report that can be generated is a blue ribbon report. In the blue ribbon report, the instructor can view the number of topics that have been passed by all of the students in a particular class or grade. Not only can the teacher see how many ribbons have been earned, but also the average percent correct for all activity on Study Island by topic. From this report, student names are again hyperlinked to their individual report.

The class summary report looks just like the individual report. It includes pre-test scores, the four categories with each sub category, and post-test scores. For the class

summary, an average percent is given in each category, as well as an average MEAP score, and the number of questions answered correctly compared to the total number of questions answered. This feature could help an instructor make decisions on how successful a class has been in a particular area and allow the teacher to decide whether or not to move on with the curriculum or give more instruction.

The class comparison chart allows the teacher to compare different classes by subject. This report gives the number of individual sessions, total time spent, number of correct compared to total questions answered, and an average percent.

All of these reports have special parameters that can be selected by the teacher. As reports are generated, the instructor first chooses whether to report information by subject, subject broken down by topic, subject broken down by reporting category, by just topic, by just subject, by program, or by all activity. The next choice is to select a grade or a specific class. Then a choice is made about the curriculum, which is connected to the MEAP scores. For example, a teacher can choose the 7th Grade GLCE Mastery (8th Grade MEAP). These parameters allow the teacher to customize reports to meet their needs. All these reports can also be filtered by date, to allow a teacher to focus on a particular time of the year if they choose to do so. There are other reports that can be generated, but they were not used in this study.

Procedure

Pre-MEAP Activities

At the beginning of the school year, all of the participants, both in the experimental and comparison groups, were introduced to Study Island. It typically takes about two days to get everybody logged in and to work out any glitches that may occur

when logging in. For example, a few students were new to the district and needed to be given access to the site by the building principal. Some students had forgotten their password, since they had changed it so it was not recorded on the password list that was given to the teacher/researcher.

In September 2009, a pre-test was administered to all students using the Study Island (SI) program. Completing the pre-test is required in order for students to participate in the program components that follow. The Study Island pre-test also gave the instructor/researcher an idea of students' prior knowledge when entering seventh grade. Following the pre-test, students engaged in some review of the mathematics content material that they were to have learned the previous academic year in preparation for the MEAP test. Study Island was used in all classes during a supplemental lesson in September when the concept of frequency was introduced. All classes were given notes and worked through guided problems with teacher direction. Students were then given a six-problem worksheet generated on Study Island as homework. The students did not receive any more SI instruction until November. All students took the MEAP test in October. The mathematics results from this 2009 test served as baseline data for the study.

Use of Study Island with the Experimental Group

The General Goal. Each week the students in the experimental group spent one whole or partial class period in the computer lab or classroom using SI to work on problems related to one of the GLCE subcategories. The problems the students worked on were selected by the instructor to correlate with material being taught in class.

Students' comprehension was monitored on the website, which reports the number of problems each student completed and the percent he or she got correct.

After the Study Island results were reviewed, subsequent instruction was given in the classroom that included both guided and independent practice. The goal was to get 75% of the class to achieve 60 % or greater on the assignment. If the goal was not met, the instructor would look at the types of problems that were missed and include them in classroom instruction as examples or in notes students were given. Following instruction, problems from the Study Island program were given to the students in a worksheet or as an internet-based homework assignment. Students who were not able to access the internet at home were given time to access the internet at school during their homeroom period. Whenever possible, the subcategories that students were practicing online became part of the units that were being taught concurrently in class.

Prior to using SI during instruction, the researcher/instructor created a homepage in SI for the students so that when they logged on to the program, the first thing they saw was a message from the teacher or an assignment that they had to do. The homepage also included a link to the National Library of Virtual Manipulative (nlvm.usu.edu/) for the students to use as a tool for thinking through problems and also as a quick way of getting to the site when technology was integrated into the classroom lessons.

Study Island Survivor. The homepage was used as the homework distribution site for a SI Survivor Game. The experimental group (1st hour class) was introduced to the game in early November, shortly after the MEAP test was given. The game was based on the Survivor reality show (Probst, 2010) seen on television. The class was divided into

two teams, and the teams choose names: PLUCK-LUCK-LOU and The Wolverines. A bulletin board was created with pictures of the students on each team. Each Monday, a Study Island assignment with ten problems was added to the homepage and students had until Thursday evening to complete it. The assignments were coordinated with the topics being taught in class whenever possible. The person with highest percent correct on the assignment gained immunity. On Friday, a Math Challenge would occur. The challenges were competitive team relays in which students had to work problems about basic math facts early on, with concepts becoming more advanced throughout the year. For example, the first challenge was a multiplication relay. A student would go to the board, grab a card, and write the problem with the answer on the board. Each team had 20 cards. The first team to answer all 20 questions correctly won the challenge. If a student answered a question incorrectly, it went back into the pile. The team that lost the challenge had to vote one person off the island; they used their notebook to do this. Notebooks were checked and votes tabulated. The castaway's picture was then moved to the water on the bulletin board.

In order to keep students interested in the game, a small monetary prize was instituted for the winner, and the instructor gave homework credit for each assignment and allowed the castaways an opportunity to get back on the island. Castaways could get back on the island if they were able to get two consecutive blue ribbons in Study Island. To get back on the island later in the year, the blue ribbon count increased from two in a row to three and then four. After the first challenge, it was realized that the students who needed to be in the challenges the most were the ones who were going to be eliminated

from their teams, as they had the weakest basic skills and slowed the team down. To remedy this, on the next challenge, the team that won the challenge eliminated a person on the opposing team. Now the better students were being cast into the water, but they were also able to get back on the island easier, as they were better swimmers, mathematically speaking.

Supplementing the Curriculum. Whenever a concept being taught in class had a Study Island supplemental lesson or notes, those materials were incorporated into the lesson plan and used in conjunction with instructor notes. These notes were used to discuss the topic in the experimental class before students started using classroom time on the computer or net book to work Study Island problems.

Study Island was also used to teach concepts that were not addressed well in the course textbook. For example, in December, students were taught math properties, such as inverse operations, the distributive, associative, and commutative properties, and the identity element, in all 7th grade math classes. This was a topic that the school's adopted math textbook did not do a great job of explaining, so supplemental material from Study Island was integrated into the instruction. Whenever supplemental material was needed for the curriculum, the experimental class used a computer to actually work out problems and view the explanations for the problems they missed. The other two classes just got worksheets that were created using the Study Island software and were not given an opportunity to work with the SI software during class.

In December, after using SI for two months, the students in the experimental group went to the computer lab to complete a review. The instructor had students work

on five concepts that were covered in class during the time between September and December: squares and cubes roots, frequencies, linear functions, scatter plots, and properties of arithmetic. The students did five problems from each category and recorded their success on 3x5 cards that they handed in. These cards gave the instructor an immediate hard copy of the number of problems they got correct, rather than having to access an administrator report. This collection of data also enabled the instructor to verify that the correct report was being used when the data was compared to the various report forms that he was learning to use in SI. The cards were also a ticket to leave the class at the end of the hour and used to keep the students accountable for doing the assignment. The instructor's evaluation of student work indicated that more instruction would be needed on frequencies and linear functions, but this would be done at a later date.

Remedial Work. In January, the middle school moved to a new building and teachers were given new technology. All classes now had access to the computer and internet, as the teachers received 30 netbooks for the 7th grade team to use. A co-teacher also joined the instructor's experimental class. The Learning Disabilities teacher was assigned to come in and work with a couple of students who were on her case load. This gave an opportunity to use SI in a couple different ways. Having an extra teacher in the room allowed for division of the students into smaller groups and gave them special attention in the areas that were giving them the most difficulty.

A plan was devised to have SI breakout sessions in the 1st hour classroom. Between April 15 and June 2, six breakout sessions were used to work on student remediation. A few students were earning one blue ribbon after another and not

experiencing too much difficulty with most concepts. These students became instructors for the remediation breakouts. Between the three students and two teachers, students were divided into groups based on their performance. Each group had one net book to use, as well as paper and pencil. A problem randomly selected by Study Island and from the topic that was assigned to the groups would be put up on the computer screen, and the students would work together to try and solve it. The students did five problems together and then they were able to get their own net book to log into Study Island and try to answer enough questions correctly to get a blue ribbon, while the three students and two teachers went around helping those who had questions. As the end of the year approached, these breakout sessions became more individualized and not as much group centered as topics of difficulty became more scattered. The instructor's goal for the students was to achieve a minimum of 60% in each topic regardless of whether or not they got a high enough percent for a blue ribbon.

Some of the students had struggled early in the year, so the instructor reset their number of questions answered to zero. The reason for this is because they would not be able to get a blue ribbon without having to answer a very large amount of questions. For example, if a student answered 50 questions and had 10 of those questions correct, they would have a score of 20%, but needed to have 60% correct to earn the blue ribbon. Thus, they would need to answer 50 more questions, all correctly, to get the blue ribbon. In these breakout sessions and in remediation sessions that were held at the end of the year, the goal was to achieve a blue ribbon. To motivate students to try to achieve blue

ribbons, students were given five extra credit points for each blue ribbon they received in a given marking period.

As the class got closer to having a Study Island “Survivor” the instructor started having additional remediation sessions, where students who were no longer on the island worked to earn their blue ribbons and extra credit points. A report that gave the percent of each of the 18 topics that each student had passed was used to plan the remedial work. The instructors would discuss with each student which categories they should work on for a given day. Again, the two teachers and a couple of students who had received all their blue ribbons went around and helped students when they had a question. The eight students who were still on the island were given a short SI test of 10 questions that could be from any topic. The person with the highest percent correct was safe. The two with the lowest percent correct were eliminated. This work continued for three weeks until a winner emerged. In the end, 7 students received all 20 blue ribbons, 1 student had 19, 1 student had 17, 9 students had between 10 and 15, and one student had only 3 blue ribbons.

Use of Study Island with the Comparison Group

The comparison group was able to access the Study Island program as they saw the need and when it had traditionally been used as a supplement to the curriculum in the past. The comparison group classes were introduced to Study Island as a supplemental resource to be used to aid their learning for 6 of the 18 subcategories. These subcategories included derived quantities, inversely proportional relationships, arithmetic properties, algebraic expression arithmetic, similar triangles, and frequencies. For each of

these categories, the comparison group was assigned printed worksheets that were generated using Study Island for homework, in addition to being given notes and guided practice problems during class. Some of the guided practice problems were given via an LCD projector hooked up to a computer that was logged into Study Island. Students were guided by the teacher to find the right answer from the multiple choice responses listed and were then presented with an explanation of the correct solution. These activities using SI generally took place during an approximately 20-minute block of time during a given lesson.

Post-test

A post-test using the Study Island program was given at the end of the year to all students to document changes in student knowledge related to the 18 Study Island subcategories. This post-test served as an assessment of the classroom curriculum as well as a way to document student achievement at the end of the year. The post-test was also used to predict 2010 MEAP outcomes.

Data Collection and Analysis

Data collection focused on gathering information about students' use of Study Island, including the results of their work and the total time spent on Study Island by individuals and specific classes. In addition, student MEAP scores were collected to assess the extent to which the use of Study Island improved student achievement on the statewide test. Each data source is discussed in more detail in the following.

MEAP Results

The 2009 and 2010 MEAP mathematics test results provided data related to how

each individual student performed on the test. MEAP scores are given as a number that indicates performance at Levels 1 through 4, with Level 1 being the highest category of achievement and Level 4 being the lowest. Scores also include a High, Middle, or Low notation so schools can easily determine whether a student's scores increased or decreased from the previous year. For example, a score of 2M means a student scored in the middle of the score range for a level 2 classification. Therefore, a student may score in the Level 2 range for two consecutive years, yet there still can be a variation in their score if the second year they received, for example a score of 2H.

Students' scores on the 2009 MEAP test for the experimental group were compared to the corresponding student scores in the comparison group to determine whether the two groups were comparable at the start of the study. For this analysis, an unpaired t-test was used with a 5% significance level. This, along with a comparison of corresponding student scores on the 2010 MEAP test, also enabled the instructor to see if a relationship existed between students' MEAP scores and the treatment received.

Comparisons of the changes in results from the 2009 and 2010 MEAP tests of the scores of students in the experimental and comparison groups allowed for analysis of whether there might be a relationship between the use of the SI program and changes in students' scores on the standardized test in mathematics. To analyze the scores, a student's score on the 2009 MEAP test was placed in a spreadsheet and their corresponding score from the 2010 MEAP test was placed beside it. A paired t-test with a 5% significance level was used to determine whether either group's scores changed significantly from one year to the next.

The experimental group percentage was compared to the comparison group percentage of students who improved, maintained, or declined in their performance level to determine whether the experimental group percentage was higher in any category. Data was evaluated based for MEAP performance levels, as well as for MEAP movement within the levels of achievement from one year to the next.

Results for specific Grade Level Content Expectations (GLCEs) were also analyzed. In this case, unpaired t-tests were used to determine whether there were significant differences between the comparison group and the experimental group performance in any of the four main mathematics subcategories associated with the state curriculum and the MEAP test. These tests were done to see if the treatment may have helped the experimental group in certain areas of instruction.

Study Island Records

The Study Island program tracks the time students spend online with the program and the number of standards they successfully complete, indicating both percent scores and the number of blue ribbons earned. The percent scores for pre- and post-tests are also scored and tabulated by the computer program.

The SI data was used to compare performance of students in different classes and on different mathematics topics. As was done with the MEAP data, comparisons were made between pre- and post-test data. An unpaired t-test was done on the SI post-test to see if the experimental group performed better than the comparison group. Once the t-test was completed, a comparison between the pre- and post-test scores for each group was done to see which group had the larger growth from pre- to post-test performance.

Unpaired t-tests were used for both the time spent working on the program during the year and on blue ribbons earned to determine whether there were significant differences between the groups in these areas. In addition, a regression analysis was used to see if there might be any correlation between time spent on SI or blue ribbon achievement and 2010 MEAP performance.

The SI category breakdown data and the comparable data for the 2010 MEAP results were used to determine whether there was a connection between the areas of difficulty that a student experienced with SI and areas of difficulty on the MEAP test. Group averages for performance were found for the experimental and comparison groups in each of the four GLCE categories. This data was put into a table and analyzed for possible relationships.

Finally, to determine whether the SI program provides a good prediction of MEAP performance, the SI post-test scores were compared to the 2010 MEAP results. If SI is functioning as intended, the post-test scores should be very similar to 2010 MEAP results. A cut score was determined for the post-test in order to compare it to the MEAP. The cut score was significantly higher than the actual cut score that ended up being used by the state, but was a score that was a fair mark for work performed by students throughout the year in 7th grade. Percentages of students performing above this cut score were then tabulated for both the experimental and comparison groups for both the SI posttest and the 2010 MEAP test. In addition, a regression analysis was used to determine whether there was a relationship between a student's SI post-test score and their 2010 MEAP result.

Researcher Journal

In order to document the research activities, the researcher kept a journal throughout the study. On the inside cover of the journal, a copy of the Study Island Grade Level Content Standards was pasted. This allowed the instructor to choose the strands that applied to the content being covered in class. It also allowed the instructor to make notations about performance on those tasks. The instructional goal was to get the experimental groups' (first hour math class) average at 70% or greater in each subcategory. If this was not achieved, a notation to re-teach the topic at a later time during the year was made.

The journal contained short entries pertaining to the date and the specific SI activities that the instructor conducted and students were involved in. These activities include "Survivor" challenges, topics covered, homework assigned, and time logged using SI that was not associated to student computer use (i.e., printed worksheets, etc.). The journal also includes some print outs of student scores on homework assignments as well as some class reports used to help evaluate student success. The role of the journal in this study was to help keep ideas and activities documented, to help the researcher keep track of ideas related to the methodology of this study, and to keep a record of time using SI materials that would not have been recorded by the computer program.

Chapter 4: Results

This chapter will focus on results associated with the data collected. Information on the students' MEAP performance, Study Island results, and a comparison between Study Island and MEAP testing will be discussed. These results will be used to determine whether the sustained use of Study Island in the study had an impact on the MEAP standardized test results, whether the SI data is a good predictor of MEAP performance, and what areas of the HMS 7th grade mathematics curriculum may need improvement.

Data Summary

Summaries of the data collected for the study, including all of the participating students and each subgroup of students is presented in the following tables. For all students combined, the data indicates that the 2009 MEAP results were very similar to the 2010 MEAP results, with a mean score on the 2009 test of 53.69% and a 2010 mean test score of 53.43% (see Table 1). The median value of the 2009 test was 53% compared to a median value of 55% for the 2010 MEAP test. The standard deviation for the 2009 MEAP was 14.81%, which is similar to the 2010 MEAP standard deviation of 12.75%.

Table 1

Participant Summary Data for MEAP and Study Island Tests

Whole Group Data				
	09 MEAP	10 MEAP	SI Pre	SI Post
Mean	53.69	53.43	42.82	71.49
Q1	41.5	43	30.5	60
Median	53	55	41	73
Q3	66	64	51	83
St. Dev.	14.81	12.75	17.97	13.48

The SI pre-test mean score was 42.82% and the SI post-test score was 71.49%. This difference indicates growth for the participants throughout the school year.

The data in Table 2 and Table 3 break down the data for the experimental and comparison groups for the 2009 and 2010 MEAP test, as well as the SI pre- and post-tests. The experimental group scored an average of 57.21% on the 2010 MEAP compared to 50.07% on the 2009 test (Table 2). The comparison group scored about the same on the test in 2009 and 2010, with averages of 51.54% and 51.91%, respectively. Overall, the data shows that the comparison groups scored were consistent for both tests and that the experimental group had a slight decline in mean score on the 2010 MEAP. On the Study Island test, both groups improved from pre- to post-test.

Table 2

Experimental Group Summary Data for MEAP and Study Island Tests

	09 MEAP	10 MEAP	SI Pre	SI Post
Mean	59.07	57.21	51.36	76.07
Q1	51	47	33	63
Median	63	58	46.5	78.5
Q3	70	65	76	87
St. Dev.	15.86	13.67	21.1	11.47

Table 3

Comparison Group Summary Data for MEAP and Study Island Tests

	09 MEAP	10 MEAP	SI Pre	SI Post
Mean	51.54	51.91	39.4	69.66
Q1	40	43	30	60
Median	51	53	40	71
Q3	64	59	50	80
St. Dev.	14.03	12.23	15.6	13.93

Comparison of Groups at Start of Study

To determine whether the experimental and comparison groups were comparable in their mathematical abilities at the start of the study, the SI pre-test results and 2009 MEAP scores for the two groups were compared.

An unpaired t-test was done for the 2009 MEAP mathematics results for the comparison and experimental groups. The unpaired test was used because two different groups of students' scores on the same test were being compared. The mean score for the experimental group (reported as a percent) was 59.07 with a standard deviation of 15.99 and the mean score for the comparison group was 51.54 with a standard deviation of 14.03 (see Appendix B for data). The t-test resulted in a p-value of 0.0678; this difference is considered to be not statistically significant at a 5% significance level. The SI pre-test was also compared using an unpaired t-test. In this case, the p-value was 0.0349, which is statistically significant. The experimental group had a mean score of 51.36, while the comparison group had a mean score of 39.4.

This comparison indicates that the two groups' MEAP scores were not significantly different at the start of the study, but the SI pre-test did indicate some differences between the groups. This initial difference will be considered when comparing student results at the conclusion of the study.

MEAP Test Results

Changes in MEAP Scores

The results of the 2010 MEAP test—administered at the conclusion of the study—for the two groups (see Appendix B for data) were compared using an unpaired t-

test. The mean score for the experimental group was 57.21% with a standard deviation of 13.66 and the mean score for the comparison group was 51.91% with a standard deviation of 12.23. The t-test resulted in a p-value of 0.110, which is considered to be not statistically significant. This test was used to see if there were significant differences in how well the experimental group did on the 2010 MEAP test in relationship to the comparison group. Since neither this t-test nor the one performed on the 2009 MEAP test data indicated significant differences between the groups, it cannot be determined from this test alone whether the use of Study Island affected students' MEAP results.

For the experimental group, a paired t-test was used to determine whether there were significant differences in the 2009 and 2010 MEAP results. The paired t-test allowed for student A's test result in 2009 to be paired with his or her 2010 result. The paired test gave a p-value of 0.3068, which is not statistically significant. This means that it cannot be determined that the use of Study Island affected student's MEAP scores. For the comparison group, a paired t-test between the 2009 and 2010 MEAP results gave a p-value of 0.4192, which is also not statistically significant. Thus, there was no evidence to suggest that the use of Study Island affected student's achievements on the standardized testing. It does not rule out the possibility that past exposure to SI may have contributed to 2009 test scores, but for this study, no clear evidence is present to support the use of SI from looking at only these results.

MEAP Scores by Level

The MEAP test results also included a report of scores by performance level. As can be seen in Table 4, the percent of students in the study that obtained a Level 1 score

in 2010 was 53%, Level 2 scores equaled 43%, Level 3 was equal to 4%, and no students performed at Level 4. These scores account for 96% of students scoring an acceptable mark on the 2010 MEAP test, which is considered Level 2 or above. In 2009, when the same students were tested, students scoring at the Level 1 mark accounted for 61%, Level 2 student scores were equal to 33%, Level 3 scores accounted for 6%, and there were no Level 4 scores (see Table 5). The 2009 test saw 94% of the study group scoring acceptable marks on the MEAP test.

Table 4

Hancock Middle School 2010 Mathematics MEAP Results by Performance Level

2010 MEAP Scores						
	All Students		Experimental Group		Comparison Group	
	Number	Percent	Number	Percent	Number	Percent
Level 1	26	53%	8	57%	18	51%
Level 2	21	43%	6	43%	15	43%
Level 3	2	4%	0	0%	2	6%
Level 4	0	0%	0	0%	0	0%

Table 5

Hancock Middle School 2009 Mathematics MEAP Results by Performance Level

2009 MEAP Scores						
	All Students		Experimental Group		Comparison Group	
	Number	Percent	Number	Percent	Number	Percent
Level 1	30	61%	11	79%	19	54%
Level 2	16	33%	2	14%	14	40%
Level 3	3	6%	1	7%	2	6%
Level 4	0	0%	0	0%	0	0%

If one looks closely at the results, it can be noted that in 2010, there were less students scoring at Level 1 than in 2009, with a corresponding increase in those scoring at

Level 2. Overall, however, the percentage scoring at an acceptable level went up, which is the goal of every school in Michigan. It should also be noted that in 2009, 7% of the experimental group scored at the Level 3 range compared to 0% at Level 3 on the 2010 MEAP, on which all students scored at a successful level. The comparison group had 6% of its students scoring at Level 3 for the 2009 MEAP test and in 2010 they still had 6% of the students scoring at Level 3. Although the student numbers are very small, and need to be interpreted with caution, these statistics indicate that SI may have been a factor in helping the lower performing students.

The MEAP results presented to school districts, track progress from year to year by including for each student the distinction of improvement, maintenance, or decline. The students in the experimental group had 11 out of 14 students at Level 1 on the 2009 MEAP test. Of the 11 students, 7 stayed at Level 1, while 4 students dropped into the Level 2 category on the 2010 MEAP (see Table 6). There were two students in the Level 2 classification in 2009; one of them moved into the Level 1 category and one stayed at Level 2. The only student in the Level 3 category went up. This resulted in the experimental group having a total of two students (12%) that increased their scoring from 2009 to 2010, eight students (68%) that stayed in the same classification, and 4 students that moved down, for a 20% level decrease.

The comparison group had a total of 35 students in the study. Of these, 25 (72%) stayed at the same scoring level, 6 (17%) dropped to a lower level, and 4 (11%) moved up to the next level (see Table 7). Nineteen out of 35 students in the comparison group were recorded as being at Level 1 on the 2009 MEAP test, while in 2010, 15 (43%) of

Table 6

Hancock Middle School 2010 Mathematics MEAP Level Movement for Experimental Group

	Level in 2009	Level Change from 2009 to 2010		
	Number	Stayed	Increased	Decreased
Level 1	11	7	0	4
Level 2	2	1	1	0
Level 3	1	0	1	0
Level 4	0	0	0	0

those students remained at this level, with the other 4 (11%) students moving down to Level 2. Two (6%) students from the comparison group scored at Level 3 in 2009, but both students improved their scores for 2010. Because two students from Level 2 in 2009 dropped down to Level 3 in 2010, no net gain was reported for Level 3.

The difference to be noted between the comparison and experimental group is how students moved between levels. The experimental group did have a higher percentage of students falling from Level 1 (36%); however, the students that were at Levels 2 and 3 improved or maintained to leave no one in the group at a scoring level less than 2. The comparison group had two students in Level 3 that improved their scores on

Table 7

Hancock Middle School 2010 Mathematics MEAP Movement for Comparison Group

	Level in 2009	Level Change		
	Number	Stayed	Increased	Decreased
Level 1	19	15	0	4
Level 2	14	10	2	2
Level 3	2	0	2	0
Level 4	0	0	0	0

the 2010 MEAP test enough to leave Level 3, but two other students from Level 2 fell to replace them with Level 3 scores. These findings suggest that SI may have an impact for the middle to low level performers.

In addition to showing scores of Level 1 to 4, the MEAP results also indicate level subcategories of high (H), medium (M), and low (L). Thus, when taking these subcategories into consideration, a student can improve or decline without leaving a level. For instance, a student may have scored a Level 2M in 2009 and be noted as declined in 2010 while still scoring at a Level 2 because they scored a 2L. On the 2010 MEAP test, overall the experimental group had 50% of the students decline in score from their 2009 MEAP results, 28% improved, and 22% maintained their results from the previous years test when taking level subcategories into consideration (see Table 8). The comparison group had 40% decline from their 2009 scores on the 2010 MEAP test, while 26% improved, and 34% maintained score results from the previous year.

Table 8

Hancock Middle School 2009 to 2010 Mathematics MEAP Result Changes (Including Changes in Level Subcategories)

	All Students		Experimental Group		Comparison Group	
	Number	Percent	Number	Percent	Number	Percent
Improved	13	27%	4	28%	9	26%
Maintained	15	31%	3	22%	12	34%
Declined	21	42%	7	50%	14	40%

Looking at the experimental group by level and including the subcategory designations, 55% of the students who scored a Level 1 in 2009 declined in 2010, 18%

improved, and 27% maintained. In the comparison group, 40% of the Level 1 students declined, 20% improved, and 40% maintained (see Table 9). Level 2 results were also analyzed. The experimental group had 50% decline, 50% improve, and 0% maintain, while the comparison group had 58% decline, 21% improve, and 21% maintain (see Table 10). The Level 3 scores in the experimental group saw a 100% of students improve, while the comparison group saw 100% improve (see Table 11).

Table 9

Hancock Middle School 2010 Mathematics MEAP Results Performance of Level 1 Students (Including Changes in Level Subcategories)

	All Students		Experimental Group		Comparison Group	
	Number	Percent	Number	Percent	Number	Percent
Improved	6	19%	2	18%	4	20%
Maintained	11	36%	3	27%	8	40%
Declined	14	45%	6	55%	8	40%

Table 10

Hancock Middle School 2010 Mathematics MEAP Results Performance of Level 2 Students (Including Changes in Level Subcategories)

	All Students		Experimental Group		Comparison Group	
	Number	Percent	Number	Percent	Number	Percent
Improved	4	25%	1	50%	3	21%
Maintained	3	19%	0	0%	3	21%
Declined	9	56%	1	50%	8	58%

Table 11

*Hancock Middle School 2010 Mathematics MEAP Results Performance of Level 3**Students (Including Changes in Level Subcategories)*

	All Students		Experimental Group		Comparison Group	
	Number	Percent	Number	Percent	Number	Percent
Improved	3	100%	1	100%	2	100%
Maintained	0	0%	0	0%	0	0%
Declined	0	0%	0	0%	0	0%

The data in Tables 9, 10 and 11 show that the area in which the experimental group had the biggest decline was in Level 1 performance scores from one year to the next, with 55% declining and only 45% improving or maintaining, compared to 60% improving or maintaining for the comparison group. This could be related to many factors, but it appears that those factors did not affect the comparison group in the same way. It seems as though the 7th grade curriculum may have allowed the comparison group Level 1 students to sustain their level of achievement or even improve, whereas the experimental group's treatment with SI did not seem to give the students any additional advantage.

Grade Level Content Standard Breakdown

Results for question subsets related to each of the four GLCE categories—Algebra, Number & Operations, Geometry, and Data and Probability—for the 2010 MEAP results were compared to look for significant differences that might indicate a weakness in the 7th grade mathematics curriculum at Hancock Middle School and to see if the treatment of Study Island helped students in the experimental group score better in

any of the categories, which might imply that the SI intervention helped improve scores in some GLCE categories.

The only categories in which differences between the two groups were noted were Data and Probability and Algebra. In the Data and Probability category, the experimental group had a mean score of 83% with a standard deviation of 9.42 and the comparison group had a mean score of 61.82% with a standard deviation of 20.01. The unpaired t-test resulted in a p-value of 0.00023, which is considered to be statistically significant.

Because of the large difference in standard deviations, this significance is difficult to interpret, however. In Algebra, the experimental group had a mean score of 55.64 with a standard deviation of 17.21 and the comparison group had a mean score of 46.97 with a standard deviation of 11.33, which resulted in a p-value of .0451, indicating that this difference was statistically significant, as well; however, the standard deviations were also large, making the results difficult to interpret. These results indicate that Study Island may have benefitted the experimental group by giving students in this group who needed extra work in the above categories sufficient time to work on deficiencies in these categories. The comparison group did not get exposure to the type of questions used by Study Island that were similar to ones used on the 2010 MEAP test, which may have accounted for their lower scores (see Table 12).

Table 12

*Hancock Middle School 2010 Mathematics MEAP Results By Grade Level Content**Standard*

	All Students	Experimental Group	Comparison Group
GLCE	Average Percent	Average Percent	Average Percent
Number & Operation	61%	62%	61%
Algebra	50%	56%	47%
Geometry	41%	39%	42%
Data and Probability	68%	83%	62%

Study Island Data

The Study Island statistics focused on time spent working with the program and blue ribbon performance. Students in the experimental group spent an average of 386 minutes on the program, while those in the comparison group spent an average of 122 minutes. This data is a representation of time that was recorded by the SI program. Along with the recorded time, the researcher journal documented 170 minutes of additional time for the experimental group that was not recorded by the SI software. This time was accumulated by going over notes in class that were from the SI program, assigning worksheets that were generated from SI, and time spent playing a classroom game on SI. The comparison group also had some non-recorded time spent on worksheets and notes. This time amounted to about 80 minutes. It was hard to come up with an accurate amount of time because when students worked on worksheets they completed them at various rates. An estimate of time had to be made to account for student effort on the activities.

An unpaired t-test was done on only the time that was logged on SI and it indicated that there was a significant difference between the time spent on Study Island

for the two groups, with a p-value less than 0.001. This information was not surprising. Although a few students in the comparison group spent significant time on Study Island, the majority of the students did not use the website as much as the school would like them to without some guidance from an instructor. When the total time was added in from the researcher's journal, the comparison group did average over 3 hours of time with the program as they had a mean of 210 minutes, while the experimental group averaged over 9 hrs of time with the SI program with a mean of 579 minutes. When looking at the numbers, it seems as though the amount of time spent on SI would help the experimental group maintain or improve their 2010 MEAP score.

Study Island Blue ribbons are awarded to students when they successfully complete a GLCE subcategory. There are 18 possible blue ribbons awarded for passing GLCEs, along with one awarded to everyone for taking the pre-test and one available for the post-test, for a total of 20 blue ribbons. Data was tabulated as a percent of blue ribbons earned. The experimental group earned an average of 78% of possible blue ribbons and the comparison group earned an average of 17%. An unpaired t-test resulted in a p-value of less than 0.001, which is considered to be statistically significant.

The time and blue ribbon results together suggest that because they spent more time on SI and were successful in passing more GLCEs on SI, the experimental group would be better able to maintain or improve scores on the 2010 MEAP test. One would believe that students in the experimental group should have at least performed better on the SI post-test due to the amount of time and success they had on SI in relation to the comparison group.

To see if the use of SI did, in fact, help students in the experimental group to perform better, an unpaired t-test was done on the SI post-test results. The experimental group did perform better with a mean score of 76% compared to the mean of 70% for the comparison group. The p-value was 0.05392, however, which is not quite significant at a 5% significance level. Recall that on the SI pre-test, the experimental group also performed better, with an average of 51.36% compared to 39.4% for the comparison group; thus, the post-test difference may be a result of initial differences between the groups. Also note that despite having lower pre- and post-test averages, the comparison group had the larger growth from pre- to post-test with a 30.6% improvement compared to 24.64% for the experimental group. Thus, an argument could be made that the comparison group outperformed the experimental group. However, looking at the empirical data it does appear that SI helped the experimental group's performance, as well.

Although blue ribbon success and time spent on SI seems to have shown some benefit to student success on the SI post-test, the question remains: does SI post-test success translate into an improved MEAP score?

MEAP / Study Island Correlation

Study Island Post-test and 2010 MEAP Comparison

For the 2010 MEAP test, the cut off for a Level 2 low achievement was about 36.7% of questions correct. To test whether or not the SI post-test was a good predictor of MEAP success, a cut score of 60% on SI was used to predict whether or not a student would be successful on the MEAP. A score 60% or better was passing and less than 60%

was not passing. Comparing the MEAP cut scores to the Study Island post-test scores, 41 out of 49 students or about 84% of the students in this study scored 60% or better on their SI post-test, which should correlate to about an 84% proficiency rate in mathematics on the 2010 MEAP test. In fact, Hancock Middle School scored 92% proficient on the Fall 2010 8th Grade MEAP Test. A little over 50% of the 65 kids tested (all of whom are not in this study) scored at a Level 1 and 42% of the students tested scored at a Level 2. Although it would appear that the SI post-test was not completely accurate at predicting the MEAP test results, the actual MEAP scores were higher than predicted, so SI might be useful as a predictor of school success. If all students who took the MEAP in 2010 would have participated in this study, the prediction would have been around 92% based on school data.

Looking at the two groups separately, the experimental group had 100% of the 14 students pass the SI post-test and 100% of the group also passed the 2010 MEAP test. The comparison group had 77% of the 35 students pass the SI post-test and only 2 out of the 35 students were not successful on the 2010 MEAP test, for a 94% success rate by the comparison group. In this case, the group actually performed better on the MEAP test than the SI post-test would have predicted. Using the cut score of 36.7% that was used on the 2010 MEAP test would have predicted 100% of the 49 students involved in this study receiving a passing mark on the test, compared to the actual result of 96% of the 49 students in this study.

Figure 3 shows the relationship between SI post-test scores and the 2010 MEAP results. All 49 students' scores have been included in the graph. A linear regression

analysis was performed for the data. The results of the regression analysis showed a positive correlation between the two categories, with a correlation coefficient of 0.68. This indicates that the strength of prediction for SI post-test to MEAP is moderately strong.

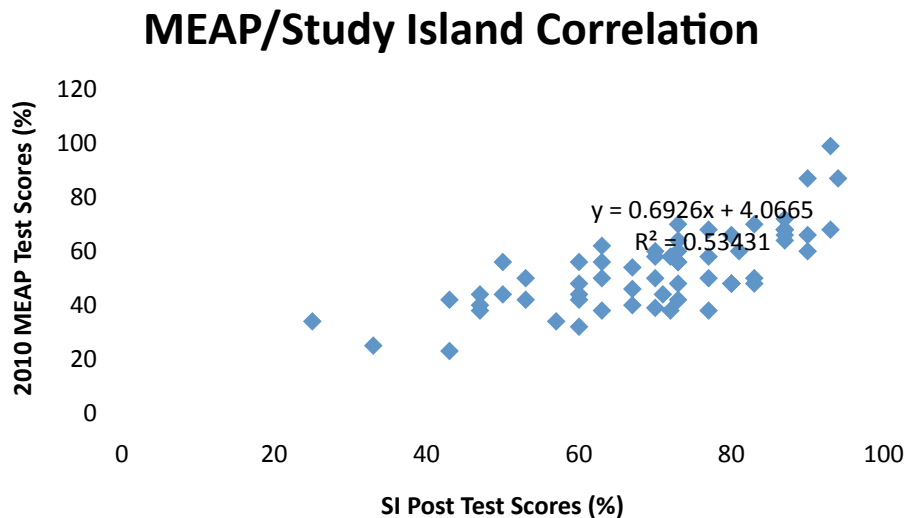


Figure 3. Relationship between 2010 MEAP Scores and Study Island Post-test Scores

Results for question subsets related to each of the four GLCE categories—Algebra, Number & Operations, Geometry, and Data and Probability—were compared between the Study Island progress data and the 2010 MEAP results. In the Study Island data, there was a noticeable difference in the average scores between the experimental group and the comparison group. As can be seen in Table 13, in the Number and Operations category the comparison group scored 57% and the experimental group scored 76% on the average. Algebra saw a mean score of 56% for the comparison group and a 64% average for the experimental group. The experimental group had a 63% mean score in Geometry, compared to a 50% average for the comparison group. For Data and

Probability, the experimental group had a mean score of 67% and the comparison group posted an average of 44%. Thus, the experimental group outscored the comparison group in each subcategory. These results may be somewhat misleading, however, because many of the comparison group participants did not post scores in some or all of the categories because they did not put in much time using SI and their scores were figured into the data as scores of zero. It may be that these results do not necessarily demonstrate a weakness in the curriculum, but a need for teachers to influence students to use SI to get the necessary benefits of its use.

Table 13

Hancock Middle School Study Island and MEAP Results By Grade Level Content

Standard

GLCE	All Students		Experimental Group		Comparison Group	
	SI Average Percent	MEAP Average Percent	SI Average Percent	MEAP Average Percent	SI Average Percent	MEAP Average Percent
Number & Operation	63%	61%	76%	62%	57%	61%
Algebra	60%	50%	64%	56%	56%	47%
Geometry	54%	41%	63%	39%	50%	42%
Data and Probability	53%	68%	67%	83%	44%	62%

Is the SI post-test a good predictor for schools to use in estimating future MEAP results? Based on the information presented by the study with the 60% cutoff, it appears that it is. With the overall success on the SI post-test at 84% compared to the actual 92% scored by HMS on the 2010 MEAP, it would be reasonable to assume that a student with a score of at least 60% on the SI post-test would receive a passing mark on the MEAP. If

one considers the linear regression analysis done above that showed a moderate positive correlation, along with results compared to the state cutoff, it appears that the SI posttest was a strong predictor of performance, as it was within 4% of the 96% of study group students scoring at a successful performance level on the MEAP.

Time Spent on Study Island and 2010 MEAP Score Analysis

Earlier in this chapter a comparison of time spent using Study Island was made between the experimental group and the comparison group. Here an analysis of the relationship between time spent on SI and 2010 MEAP results is presented. A linear regression was done on the data to compare the time spent on SI to the 2010 MEAP results. The findings demonstrated that there was no correlation between time and MEAP results for either the experimental group ($r = -0.045$) or for the comparison group ($r = 0.068$). The data earlier in the chapter pointed to the great significance of time that the experimental group had over the comparison group. It also pointed to a slightly significant edge in performance on the SI post-test for the experimental group. This analysis, however, showed no relationship between time spent on SI and MEAP score, meaning the benefit of time may have helped on the SI post-test but did not seem to lead to any advantage on the MEAP test.

Study Island Blue Ribbon Percent and 2010 MEAP Score Analysis

Blue ribbon achievement was also looked at previously in this chapter. The results did demonstrate that the difference in amount of ribbons achieved between the experimental group and comparison group was statistically significant between the two groups. When the ribbon performance rate was compared to the 2010 MEAP results, a

slight positive correlation was demonstrated. When the data was plotted a weak positive correlation was noticed (see Figure 4). When the linear regression was completed it demonstrated a positive correlation, with an r-value of 0.332. This indicates that a higher rate of blue ribbon percentage is weakly correlated to a higher performance on the MEAP test.

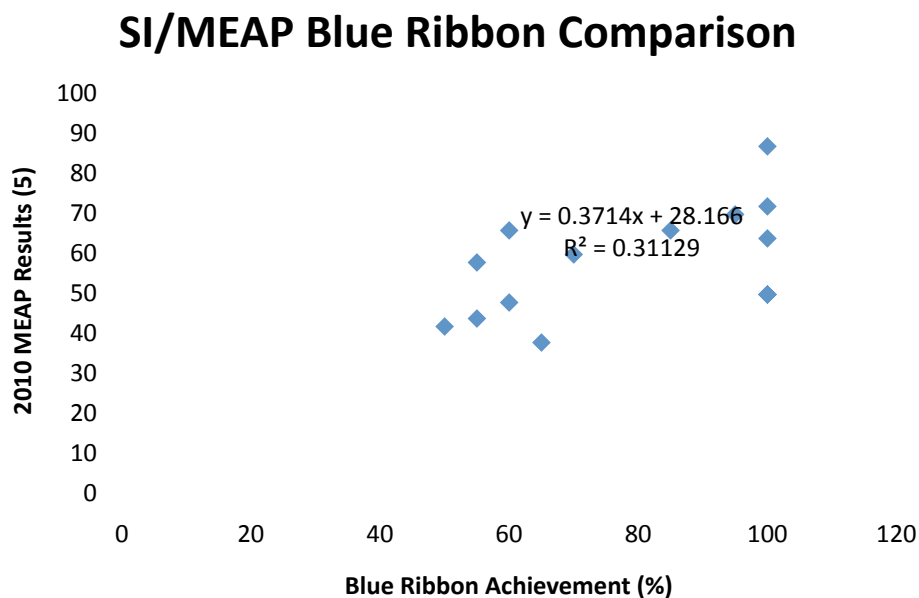


Figure 4. Relationship between Experimental Group 2010 MEAP Scores and Study Island Blue Ribbon Performance

Assessment of Curriculum

In an attempt to analyze the effectiveness of the 7th grade math curriculum at Hancock Middle School, the data for the four test subcategories were looked at. The experimental group demonstrated more success on the MEAP in three of the four categories and outperformed the comparison group in all four subcategories on the SI post-test (see Table 10). This data indicates that the use of SI with the experimental group

may have strengthened the regular classroom curriculum. The one category in which the comparison group scored better on the MEAP test was the Geometry category, although the results were very close between the two groups. The SI results would have predicted that the experimental group would have scored better. It may be the case that even though both groups had the same curriculum, the supplement of SI may have detracted from the curriculum, and the comparison group may have benefitted from a more focused textbook approach which allowed them to perform better on the geometry questions. However, the Geometry scores for the comparison group were only 42%, leading one to believe that Hancock Middle School may want to address the overall effectiveness of the curriculum in this area.

Chapter 5: Conclusions and Implications

Based on the data, there are mixed results in support of Study Island and its real benefits for supporting student learning. If the SI treatment was beneficial to the experimental group, then one would expect noticeable differences in results on the 2010 MEAP test; however, there was no considerable difference in the levels at which the experimental students scored from the 2009 to the 2010 test, nor between the 2010 scores of the experimental and comparison groups. The comparison group also did not have any exceptional growth from one year to the next; in fact, the two groups were quite similar in performance. The comparison group actually had a greater percentage of students who improved or maintained their scores from year to year, with 60% maintaining or improving, while the experimental group had a combined 50% of students who maintained or improved their score. When the scores were broken down into MEAP levels of achievement, the students at Level 1 in the comparison group in 2009 performed better by having 60% of the group maintaining or improving, compared to only 45% of the experimental group.

The experimental group put in a significant amount of time on SI as part of their treatment in relationship to the comparison group and earned significantly more blue ribbons, indicating that they completed more of the GLCEs in the program. There was no demonstrated correlation between time spent on SI and the 2010 MEAP results, however, meaning that the more time spent on SI did not give the students a better chance of earning higher scores on the MEAP test. Also, there was only a weak positive correlation

between Blue Ribbon achievement and MEAP results, meaning that blue ribbon performance does not necessarily correlate to higher performances on the MEAP test.

In the linear regression analysis between SI post-test data and MEAP performance, a moderate positive relationship was noted. The correlation was not strong, but evidence suggests that SI was somewhat effective in predicting MEAP performance. The results that do support the use of Study Island were noticed when the MEAP GLCE subcategories were evaluated. This category break down data from the MEAP suggested that the students in the experimental group benefitted from the use of SI on the Data and Probability and the Algebra portions of the MEAP, as unpaired t-tests for both categories showed statistically significant differences that pointed toward SI benefits.

The most important question for administrators is associated with state standardized testing. Will this program improve MEAP results? The sample size used in this study was relatively small (less than 50 students) so the results may be a bit misleading, but for Hancock Middle School in 2010, scores did improve on the MEAP test. The experimental group, however, did not improve more than the comparison group. Though the number of students who scored in the Level 2 or 3 category on the 2009 MEAP test in the experimental group was small (3 students), improvement on the 2010 MEAP test was documented as two of the three students jumped up a level and one maintained, and all scored at what is considered an acceptable level for success on the standardized test. Level 1 scores dropped in 2010, but Level 2 scores improved for a 2% increase over the previous year. Furthermore, the use of the SI post-test indicated that the school should score around 84% on the MEAP test and HMS scored at 92% on the state

standardized test, suggesting that the SI post-test is a fair predictor for future performance on the MEAP. These results, along with the regression analysis that showed a moderate positive correlation between the SI post-test results and the MEAP test results, indicate that the SI posttest was a good predictor as it was within 4% of the study group's MEAP score of 96% proficient.

When trying to answer the question of whether or not Study Island can give an idea of curricular success, my attention was drawn to the pre- and post-test data from SI. The experimental group performed better on both the pre- and post-test, but the statistic that might best demonstrate the strength of the curriculum is the growth from the pre- to post-test. The experimental group had a growth of 24.64%, while the comparison group had a growth of 30.6%, which gives credence to the strength of the 7th grade math curriculum, as Study Island was only used sparingly as a supplement to the curriculum with the comparison group.

The GLCEs also helped highlight some areas of concern in the curriculum. The experimental group posted an average of 83% compared to the comparison's group average of about 61% in the category of Data and Probability, which indicates that SI may have done a better job preparing students for the MEAP test than the curriculum in this area. Also, the area of Geometry demonstrated a weakness in the curriculum due to the fact the experimental group, even with the treatment, had an average score of 39% compared to 42% for the comparison group. This information shows that students using the curriculum solely performed better, but that the numbers achieved were still low,

indicating that there is room for improvement in this area in regards to curriculum development.

Anecdotally, it is my opinion that SI is a beneficial component to Hancock Middle School's mathematics curriculum for reasons that go beyond the data analysis in this study. Schools across the state are struggling financially, while at the same time trying to update textbooks and align them with the state curriculum. Textbooks can cost thousands of dollars that schools do not have. When faced with using out-of-date textbooks, the SI curriculum can help supplement classroom instruction as well as provide students with an alternative way to learn with the integration of computer technology. SI may not have been the end-all cure for the students in this study but it is apparent that it did benefit students in at least some ways. Additionally, I feel that it helped those students in the mid to lower levels of achievement, as it gave them a means to get the extra practice and time they needed on challenging concepts that they may not have fully understood in class.

Finally, SI may not have been introduced as often in the comparison group, but my use of SI in the experimental group very possibly affected the way I taught. From the extensive use of SI during the study, I have become far more familiar with the grade level content expectations and more aware of what has been covered and what needs to be covered in class with the students, as well as areas that need to be better addressed in our curriculum. For that reason alone, SI has been very helpful in the education of students in 7th grade math at Hancock Middle School.

Limitations of the Study

The main limitation relates to the number of students included in the study. Hancock Middle School is a relatively small school, so the number of kids in the experimental group was small and I believe that having more kids in the experimental group would lead to better and possibly different results. This was a particular problem with the t-tests that were performed as part of the data analysis, since larger numbers would have led to more valid results. Also in regards to the students, I had to use the first hour class as the experimental group because of the availability of the computers. Had I been able to use a different class with fewer students that scored in the Level 1 range on the 2009 MEAP test, it is possible that results would have been different. When using a small group of students for research such as this, it is quite possible that the composition of the group would change the outcome of the results.

Another limitation in a study such as this with the internet being the main component is the technology. The limitations include having enough computers to allow all students access, the availability of computers during the time that a class meets, and having enough time to set up during a 45-50 minute class. In the beginning when we were in our old building, there were times when the computer lab was signed out for the morning that I was trying to use the lab and I would have to rearrange my lesson plans to fit in the SI activities for the week. Once we moved to the new building, I had access to netbooks that I could use in my classroom and I always had a minimum of 10 of the books, which allowed availability virtually anytime I needed students to have access to

SI, but if I needed all of the students to be on SI, I still had to make arrangements to use the rest of the 7th grade netbooks that amounted to 30 units.

Technology can always be an issue. Technology is great when it works and a waste of time when it does not. Having the ability to address technological issues such as internet connections, servers, and such, as well as forgotten passwords is key in the battle over time versus technology. Many schools have an on-staff technical support to deal with those issues. The SI website provides support and we did not experience any difficulty using the site. One way around the limitation of tech support is to allow students to do materials from home and discuss during class. The biggest issue was having administrative access to be able to reset some of the students questions attempted totals to give them the opportunity to get a blue ribbon after they had put in many rounds of unsuccessful practice. Also, my familiarity with some features of SI, like password look-up were issues in the beginning.

One of the components of the SI software was the use of i-clickers. This component would create a more game-like challenge for students and give immediate feedback. For this study, i-clickers were not available, so the students had to use index cards and work as a team, which did not have the same effect and the same type of feedback for teacher and student. The i-clickers are not a necessity for use of Study Island, but they are a luxury that would only add to the learning environment for students of the entertainment era.

Suggestions

One way to build on this study would be to extend the time of the study. The study might give some interesting results if it followed students from 6th grade through 8th grade. It would also be beneficial to have a teacher from each grade collaborating on SI uses within the curriculum as a form of computer aided instruction. Another similar modification to this study would be to use the principle research ideas presented in this study using the entire 7th grade class from HMS as the experimental group. One or two comparable schools with similar demographics that do not currently use SI could comprise the comparison group. This would eliminate the teacher as a factor in the results. In this study, the teachers' instruction may have compromised the results because of influences on his instruction from use of the program in the experimental group. In this modified study, the sample size would also be larger, which could give a more accurate assessment of the influences of the SI program.

References

- Ash, J. E. (2004). The effects of computer-assisted instruction on middle school mathematics achievement. Orchard Software. Retrieved from <http://orchardsoftware.com/docs/OrchardStudyAsh.pdf>
- Bouffard, S. M. (2006). "Virtual" parent involvement: The role of the internet in parent-school communication. (Unpublished doctoral dissertation). Harvard Graduate School of Education, Cambridge, MA.
- Cotton, K. (2007). Computer-assisted Instruction. *Education Northwest*. Retrieved from <http://usc.edu.tt/faculty/fbaldeo/EDTE%20330/Computer%20Assisted%20Instruction.pdf>
- D'Arcy, C. J., Eastburn, D. M., & Bruce, B. C. (2009). How media ecologies can address diverse student needs. *College Teaching*, 57(1), 56-63.
- Gregorc, A. F. (1984). *Mind Styles*. Retrieved from <http://web.cortland.edu/andersmd/learning/Gregorc.htm>
- Kilpatrick, J., Swafford, J., & Findell, B. (Eds.) (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.
- Loong, E. & White, B. (2004). Teaching mathematics using the internet. Available at http://www.merga.net.au/documents/RR_loongwhi.pdf
- Magnolia Consulting (2008). *Study island scientific research base*. Retrieved from <http://www.studyisland.com/salesheets/SI%20Foundational%20Report%207-25-08.pdf>
- Michigan Department of Education (2008). *Michigan educational assessment program*

technical report. Office of Educational Assessment and Accountability. Retrieved from http://www.michigan.gov/documents/mde/MEAP_Technical_Manual_2008-2009_311738_7.

- Payzant, T. W., and Levin. J. (1995). Improving America's schools for children in greatest need. In *National issues in education: Elementary and Secondary Education Act*, J. F. Jennings (Ed.). Bloomington, IN and Washington, DC: Phi Delta Kappa International and the Institute for Educational Leadership.
- Probst, J. (Producer). (2011). Survivor [Television series]. New York: CBS Broadcasting.
- Skyward, Inc. (2011). School Administrative Software (Computer Software). Retrieved from <http://www.skyward.com/Page.ashx/Home>
- State of Michigan (2011). *State school aid act*. Retrieved from www.michigan.gov/documents/5I-SchoolsChoice_41452_7.pdf - 2011-09-21
- State of Michigan. (2008). *MEAP released items 7th grade*. Michigan Department of Education. Retrieved from http://www.michigan.gov/documents/mde/f08_Gr8_Math_Rel_Doc_269514_7.pdf
- State of Michigan. (2007). *MEAP released items 7th grade*. Michigan Department of Education. Retrieved from http://www.michigan.gov/documents/mde/f07_Gr8_Math_Rel_Doc_222005_7.pdf
- State of Michigan. (2006). *MEAP released items 7th grade*. Michigan Department of Education. Retrieved from

http://www.michigan.gov/documents/mde/f06_Gr8_Math_Rel_Doc_181517_7.pdf

State of Michigan (2006). *Mathematics grade level content expectations companion document*. Retrieved from

http://www.michigan.gov/documents/mde/Math_GLCES_Assessed_with_NC_Designations_v4-09_3_274952_7.pdf

Study Island (2009). [Computer software]. Available at www.studyisland.com.

Study Island (2008). Study Island Michigan statewide research data. Available at http://www.studyisland.com/salesheets/MI_Statewide_Research.pdf.

U.S. Congress (1994). Improving America's Schools Act. Available at <http://www2.ed.gov/legislation/ESEA/index.html>.

U.S. Congress (2001). No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, §115, Stat. 1425.

Webb, N. L. (1997) *Criteria for alignment of expectations and assessment in Mathematics and science education* (NISE Research Monograph No. 6). Madison, WI: University of Wisconsin Madison, National Institute for Science Education.

Wixson, K. K., Fisk, M. C., Dutro, E. & McDaniel, J. (2001). *The alignment of state standards and assessments in elementary reading*. Academia.edu. Retrieved online April 15, 2011 from http://colorado.academia.edu/ElizabethDutro/Papers/354015/The_Alignment_of_State_Standards_and_Assess-Ments_In_Elementary_Reading

Appendix A—IRB Documents

IRB Approval



Michigan Technological University

Office of Research Integrity
and Compliance

317 Admin. Building
1400 Townsend Drive
Houghton, MI 49931
906.487.2902

MEMO

TO: Dr. Shari Stockero, Cognitive & Learning Science

CC: Paul Dube, Cognitive & Learning Science

FROM: Joanne Polzien, Director Research Integrity and Compliance

DATE: October 6, 2009

SUBJECT: Approval M0490

Your application to use human subjects in research or classroom situations has been reviewed with the following determination:

Protocol #: M0490

Protocol Title: "Computer Generated Mastery Programs and Standardized Testing"

Approved Dates: October 6, 2009 through October 5, 2010

Approvals are granted for up to a one year period. You will need to request a continuation for each year of the project six weeks prior to the end date indicated above for each year of the project. The Office of Research Integrity and Compliance will make every effort to send the Principal Investigator annual reminders. However, the Principal Investigator is responsible for submitting annual Continuation Forms in advance of the expiration date for the project. It is very important that these expiration dates are not missed. Failure to submit annual review materials on time will result in the termination of this protocol.

This approval applies only for this project, and only under the conditions and procedures described in the application; if any changes are made in the protocol or conditions set forth in the application, the principal investigator must obtain a separate approval before these changes take place. The approved project will be subject to surveillance procedures requiring periodic review. This review will consist of consulting with the principal investigator and examining the appropriate project records.

Individual identification of human subjects in any publication is an invasion of privacy. Before beginning a project involving human subjects, and only if required, the principal investigator must obtain a properly executed informed consent from each subject and/or the person legally responsible for the subject. **If a consent form has been reviewed and approved it has been attached with an official date stamp on it. Only copies of the official date stamped informed consent is to be distributed to participants relating to this project. If any changes or modifications are needed regarding this form, you must first submit the revised document for review and approval prior to use.** The principal investigator must retain informed consent forms on file for at least three years after the end of the project. If a project involves a high level of risk, copies of the signed informed consent forms must be filed with the Human Subjects Committee; if this is the case, you will be notified.

This document is on file in the Office of Research Integrity and Compliance. If you have any questions, please contact me at 487-2902 or jpolzien@mtu.edu.

Participant Consent Letter

October 2009

Dear Parent/Guardian and Students,

I am currently working on research project as part of my Master's Degree in Applied Science Education program at Michigan Technological University. The purpose of this research is to determine the impact our computer generated mastery program (Study Island) has on our students' MEAP results.

In the study, the students will be working with the Study Island software that is already used at Hancock Middle School. All students will receive at least the same level of preparation for the MEAP test as they have in the past.

Dr. Shari Stockero, a Michigan Tech professor, and I are the only people who will have access to data from your child. At the conclusion of the research, the students' MEAP data and other information related to the study will be reported as a group result only; individual student work will not be identified.

There are no known risks associated with participation in this study. Participation in the study is voluntary and is determined by the student and the parent/guardian. If you or your child decide not participate in this study, it will not affect the normal services provided to your child at Hancock Middle School, nor will it affect your child's grade in seventh grade math. Non-participation will not excuse your child from using Study Island as the teachers see fit, but will eliminate the use of data associated with your child in this study. If the student and parent/guardian agrees to participate, any of these parties are free to end participation at any time.

If a student, parent/guardian chooses not wish to be a participant in this study, please notify me by returning the second page of this letter. If you would like to discuss this study with me you can reach me by email (pdube@hancock.k12.mi.us) or phone at 487-5923 ext 421. If you have any questions or concerns, please contact either me or my advisor, Dr. Shari Stockero (487-1126 or stockero@mtu.edu), at any time.

If you have any questions about your student's right as a research subject, you may contact the Michigan Technological University Institutional Review Board (IRB) by mail at 1400 Townsend Drive, Houghton, MI 49331, by phone at (908) 487-2902 or by email at jpolzien@mtu.edu.

Sincerely,

Paul Dube
Mathematics Teacher, HMS

Dr. Shari Stockero
Assistant Professor, MTU

Parental/Student Approval Form

_____ I give my permission to use my student's data in Mr. Dubes' study of Computer Based Mastery Programs and Standardized Testing.

_____ I do not give my permission to use my student's data in Mr. Dubes' study of Computer Based Mastery Programs and Standardized Testing.

Parent/guardian signature _____ **Date:** _____

_____ I give my permission to use my student's data in Mr. Dubes' study of Computer Based Mastery Programs and Standardized Testing.

_____ I do not wish to participate in Mr. Dubes' study of Computer Based Mastery Programs and Standardized Testing.

Student signature _____

Appendix B – Data Tables

Experimental Group Raw MEAP Data

MEAP Test Scores		2010 MEAP Test Results By Category			
2009	2010	Number and Operation	Algebra	Geometry	Data and Probability
43	41	36	38	25	83
26	37	36	29	25	83
64	49	57	43	25	83
53	69	71	62	75	83
53	59	64	66	25	83
77	86	79	81	100	100
64	43	43	43	13	83
79	63	71	66	25	83
38	57	79	71	25	83
70	71	57	81	50	100
62	65	71	66	38	83
70	49	71	38	25	66
51	47	57	38	38	66
77	65	71	57	63	83

Experimental Group Raw Study Island Data

SI Test Results		Time Using SI		SI Results By Category				
Pre-test	Post-test	Total Time	Time Log	Blue Ribbons	Number and Operation	Algebra	Geometry	Data and Probability
41	60	631	461	50	75	62	47	58
39	63	592	422	65	65	62	55	59
77	77	646	476	100	71	74	77	69
52	83	799	629	95	81	72	70	70
33	73	638	468	70	69	61	55	60
80	90	469	299	100	88	79	79	82
40	60	527	357	55	72	64	57	61
81	87	413	243	100	82	72	73	71
20	72	564	394	55	71	52	54	60
67	87	670	500	100	78	73	65	78
52	90	443	273	60	74	49	69	57
76	83	731	561	100	87	67	69	76
30	60	408	238	60	71	51	50	63
31	80	748	578	85	76	61	67	68

Comparison Group Raw MEAP Data

MEAP Test Scores		2010 MEAP Test Results By Category			
2009	2010	Number and Operation	Algebra	Geometry	Data and Probability
66	67	86	52	63	83
45	38	57	38	25	13
64	57	71	48	38	83
64	53	57	38	88	50
40	43	57	43	17	50
60	55	64	66	25	66
28	55	79	48	38	50
57	49	64	48	25	50
32	49				
66	67	79	57	75	66
47	47	57	38	38	66
36	37	21	38	38	66
58	61	64	57	50	83
51	67	86	57	50	83
74	67	64	71	50	83
43	57	57	52	50	83
36	31	43	24	25	33
43	43	50	38	25	66
36	33	29	38	13	50
66	45	57	38	38	50
57	55	57	48	75	50
47	41	64	33	13	50
51	37	50	33	13	50
26	43	43	33	50	66
74	65	79	62	63	66
77	86	86	71	100	100
66	59	79	48	50	66
38	37	36	43	25	33
49	59	50	52	38	100
58	69	79	57	75	66
60	47	57	57	13	33
68	55	64	52	38	66
34	41	50	38	13	66
47	47	71	43	25	33
40	55	57	38	50	83

Comparison Group Raw Study Island Data

SI Test Results		Time Using SI		SI Results By Category				
Pre-test	Post-test	Total Time	Time Log	Blue Ribbons	Number and Operation	Algebra	Geometry	Data and Probability
47	87	203	123	20	90		65	
50	70	209	129	35	68	0	60	0
16	77	189	109	20	46		60	68
42	67	214	134	25	75	100	46	70
18	50	138	58	5	32		20	
52	73	223	143	25	54		37	83
36	60	236	156	20	30	40	61	20
10	70	165	85	15	60		0	36
40	53	170	90	10	60	100	44	37
61	87	197	117	30	85		75	84
35	83	268	188	30	69	100	59	17
39	47	212	132	5	16		55	20
50	63	155	75	10	63		54	
10	77	201	121	20	64		69	57
60	93	180	100	30	77		68	58
44	70	196	116	15	67	0	35	0
50	60	152	72	5	29			
24	71	341	261	45	61	66	46	55
24	57	207	127	5	17		53	100
48	67	156	76	15	29		65	77
44	50	192	112	10	40	40	60	55
19	53	157	77	5	45		21	
40	72	242	162	40	75		62	
33	47	178	98	5	56		59	
43	87	167	87	15	80		50	27
80	94	326	246	35	86	100	70	82
50	81	148	68	10		0	42	
35	77	220	140	15	51		48	
44	90	158	78	15	70		36	17
30	73	112	32	10	90			
30	80	170	90	20	33		71	
40	63	169	89	5	59	60	57	
44	43	173	93	5	50	60	0	0
67	73	181	101	15	68	60	44	22
24	73	211	131	10	50	62	56	35