2014

The Impact of Student Centered Learning Strategies in Middle School Earth Science

Zane Jay Whittington
Louisiana State University and Agricultural and Mechanical College

THE IMPACT OF STUDENT CENTERED LEARNING STRATEGIES IN MIDDLE SCHOOL EARTH SCIENCE

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Natural Science

in

The Interdepartmental Program in Natural Sciences

by
Zane Jay Whittington
B.S., Louisiana Tech University, 2004
August 2014

This work would not have been possible without my mother, whose appreciation for education was passed to both of her sons. She tirelessly quizzed and questioned my brother and me, promoting an inquisitive spirit in both of us. She fiercely defended a son with attention deficit hyperactive disorder, which got her sideways stares and snickers. That defense helped ensure opportunities for me, and she instilled in me daily that I was just as talented as the other students that I was frequently contrasted against.

My mother fought for my education and often fought with me to make certain that I did not squander educational opportunities. My mother was my advocate, my biggest fan, and at times my counterpoint. She sacrificed personal rewards to give me experiences where I could gain valuable life lessons in areas where I was novice. As a pediatric nurse and a teacher, my mother was skilled in caring for and teaching others. She enjoyed seeing her sons succeed in school and in life, and she took full advantage of any opportunity to share with others what "her boys" were doing.

Thank you mother for your constant counsel and your direction over my life; your careful attention to my education has given me wonderful opportunities that were because of seeds you planted and sacrifices you made. My career path and decision to go into teaching is because of your encouragement. My pursuits in this program are thanks to your many investments into my life. It is with great sorrow that I will not be able to share my completed work with you; I know how happy it would have made you to share in this experience along with friends and family. The many gifts you have bestowed on me are immeasurable, but your love was by far the greatest gift I could ever receive. Thank you.

Jacquelyn "Jacquie" Lee Whittington (July 27, 1953 - June 16, 2014)

ACKNOWLEDGMENTS

Thank you to Dr. Slezak, Dr. Browne, and Dr. Cherry for their thoughtful direction and encouragement during this project. Thank you for giving me a chance to participate in a program where I did not look like I belonged. All of you have enriched my understanding of physics and what great educators look like. I am grateful for all of your help in clarifying the picture presented with my study. It was a difficult story for me to see; I would not have noticed it without considerable assistance.

Thank you to my LaMSTI cohort for making the last three years a truly enjoyable experience. I am grateful for the support I have received in the last three years in sharpening my understanding of physics. It has been a privilege to learn alongside and from this gifted group of teachers.

Thank you to my wonderful wife for enduring the long nights and endless conversations about physics and pedagogy. You are my constant and best friend and you carried me through this process during the most challenging time of my life. Your encouragement meant more than you will ever know during times when I questioned my ability to succeed in this program.

Thank you for the generous support from the National Science Foundation grant number 0928847 that allows teachers to strengthen their content and impact students in Louisiana. I am thankful for Dr. Madden, Dr. Blanchard, LaMSTI program participants and alumni, my school administration, my friends, my family, and my students. Without the help and feedback from all of you, I would not have been able to produce this work.

TABLE OF CONTENTS

ACKNOWLEDGMENTS iii

ABSTRACT v

INTRODUCTION 1
  A Teachers Conundrum 1
  Traditional Versus Student Centered Learning 2
  An Introduction to Information Graphics 5

REVIEW OF CURRENT LITERATURE 7

MATERIALS AND METHODS 15

DATA AND RESULTS 25
  Great Scholar and Regular Groups Pretest Scores on MSGI and CTSR 26
  Great Scholar and Regular Groups Results on MSGI 28
  Great Scholar and Regular Groups Results on CTSR 31
  Math Groups Pretest Scores on MSGI and CTSR 34
  Math Groups Results on MSGI 37
  Math Groups Results on CTSR 39
  Math Groups Results on MSGI by Graphical Language 40

CONCLUSIONS 46

REFERENCES 47

APPENDIX A MEASURING THE MEASURING TOOL 50

APPENDIX B PRECISION OF MEASURING 52

APPENDIX C WINGSPAN 54

APPENDIX D IRB APPROVAL 57

VITA 58

ABSTRACT

Research continues to reinforce that student centered classrooms and interactive engagement (IE) strategies, when used effectively, can produce considerable gains compared to traditional instruction methods. In this study, IE strategies, primarily modeling instruction, were compared to traditional instruction in a middle school classroom to determine if IE strategies would have an impact in two specific areas: graphing ability and science reasoning skills. Class mean scores on tests were compared over time to show that IE strategies produced greater gains in graphing and science reasoning than traditional instruction for one group. The other group did not see significant differences in graphing or science reasoning based on instructional method. Student scores were also analyzed based on math preparations, and it was noted that students enrolled in pre-algebra math classes saw statistically significant gains in graphing. Results from this study suggest that using IE strategies in middle school science may have a positive impact on graphing ability and science reasoning skills.

INTRODUCTION

A Teachers Conundrum

For as long as I have been teaching, I have heard the same teacher concerns about the students in our science classes. The science teachers at my school have a strongly held belief that the reason our students do not meet academic standards stems from their lack of understanding in the "STEM" disciplines. Students entering our school come from very diverse backgrounds; some students are entering our school from private or parochial institutions, others from home school environments, but largely our population is coming from public elementary schools in and around our community. Our students come to us with varying degrees of scientific reasoning due, in large part, to the many different elementary schools that our school serves.

The teachers often share their frustrations about students' understanding of the scientific method and how science is done. Scientific Inquiry is the first unit of instruction in each middle school grade level. We give dedicated class time each year explaining the "doing" of science: its principles, its tools, its collection, its dissemination. Despite this, the same problems persist in the middle school student body. Many of our students come in with a limited understanding of the common measurements tools used to quantify their observations and evaluations. Some students have never seen or used a mass scale or a graduated cylinder. The students that do have exposure to measurement tools do not necessarily have mastery, or even proficiency, with regard to the proper function of these basic scientific implements.

Traditional Versus Student Centered Learning

Figuring out how people learn is a complicated task, and understanding how and when people learn has been researched extensively. Teachers have always strived to find effective ways to deliver instructional materials and information to students. There has been a healthy discussion regarding when students are prepared cognitively to take on progressively more complex and sophisticated ideas or concepts. So what does it mean to learn? What is required to have knowledge of a thing? According to developmental psychologist Jean Piaget:

> Knowledge is not a copy of reality. To know an object, to know an event, is not simply to look at it and make a mental copy or image of it. To know an object is to act on it. To know is to modify, to transform the object, and to understand the process of this transformation, and as a consequence to understand the way the object is constructed (Piaget, 1964).

Learning is more than just being able to copy and repeat. Knowledge requires an ability to manipulate and transform. With so much previous research showing that learning does not happen drastically with traditional forms of instruction, it should be surprising that teachers are still using this ineffective model for science education (Halloun & Hestenes, 1985; Chickering & Gamson, 1987; Prince, 2004). In my classroom this past year, I wanted to try to utilize effective instructional methods proven by classroom research. Research has suggested that students' science reasoning skills and conceptual understanding improve more when employing interactive engagement (IE) methods in the classroom (Johnson & Lawson, 1998; Redish & al., 2000), where the learning is student-centered instead of teacher-centered.

One large-scale study done on physics courses of high school and college students (Hake, 1998) found that classroom instruction utilizing such IE strategies produced significantly different results from classrooms utilizing traditional instructional methods. The study looked

at sixty-two different courses totaling over six thousand participants and found that students' conceptual gain in IE classes dwarfed the conceptual gain of those in traditional classroom settings. The results showed clearly that instructional method can have a drastic impact on student learning and that traditional instruction was far less successful at producing an impact compared to alternative teaching strategies.

Modeling instruction is an interactive engagement (IE) teaching framework that makes use of many strategies to promote a student centered learning environment. It makes students work in groups to explore natural events, make connections between concepts, represent those connections through different modes (diagrammatic, functional, graphical, verbal), present their ideas to their peers, and ultimately explore new ways of assimilating new information into their existing mental models. Modeling instruction has been found to produce significant gains in physics conceptual understanding in high school and college physics courses (Hake, 1998), as well as in chemistry and physical science classrooms (Prince, 2004; Deakin, 2006; Crews, 2006).

Modeling instruction is starkly different from traditional instruction. Where traditional instruction places the teacher as the central figure in the classroom, modeling instruction uses the teacher as a guide. The students, and their collective, become the authority in the classroom. In traditional instruction, there is little peer-to-peer interaction because the teacher is delivering instruction, often lecturing, to the entire classroom. In modeling instruction, the students are interacting with one another far more than they are interacting with the teacher; the teacher is in the room, but often answers student questions with open-ended questioning. A common strategy in modeling has teachers redirect student questions by asking the students

what their observations tell them, instead of giving answers to students. In traditional instruction, you may see group assignments and be inclined to think that students are actively engaged in a collaborative activity because they are working in groups. Often times what the students are working on is not an activity that requires collaboration from the group, but simply requires that the questions be answered whether by individuals or the group. The answers to the questions may have even been previously given to them. They are seldom given the responsibility of creating new information, but instead are asked to regurgitate information from previous instruction at the appropriate time in the right succession at appropriate intervals. Students need only to learn the pattern of re-delivery of the given materials to get high marks in the classroom. These students are not learning science in these classrooms; students are learning how to be good students and good repeaters of information. Traditional instruction has been characterized as “sit, get, spit, forget” where the student is only responsible for memorizing material long enough to recall the facts for a test. Is this what learning is supposed to look like? Is learning just memorizing discrete facts in isolation from other discrete facts?

Modeling instruction has been supported by research that indicates that it is effective at impacting student conceptual learning, but I was not able to find where modeling instruction was effectively implemented in a middle school earth science classes. I was not certain that I would be able to adapt a modeling framework into my earth science classroom and still maintain the primary learning goals. I knew that I wanted to evaluate my instructional method and I wanted to experiment with students’ graphing ability. I knew that I wanted to use an IE strategy in the classroom. I wanted to find an IE strategy that challenged the students’

graphing abilities. I knew from research that modeling requires the student to represent information in many different ways, and that a frequent representation of science concepts was done with diagrams, charts, tables, and most importantly graphs.

An Introduction to Information Graphics

If you keep up with social media, you are aware of its influence over today's society. It is difficult to find a space in our culture where cellphones, computers, and electronics are not present. Today young people are consumed with social content thanks to applications like Twitter, Instagram, and Snapchat. Much of the content on these sites is provided in the form of pictures, diagrams, graphs, or the infamous "Meme." Social media is moving more and more towards information represented in a graphical or illustrated format. The content may have embedded information inside the image; information that requires the reader to engage, process, and interpret the meaning of the illustration.

While my middle school students are adept at finding and filtering content from their friends, I do not know if my students can effectively interpret information embedded inside of an illustration, a diagram, a table, or a graph. My students look at pictures of friends all day, but when I present them with a graphical representation of data that has been collected they are puzzled and confused. I would like to believe that my students have all just played a clever prank on their teacher, but I am presented with this same problem year after year in my science classroom. While my students spend considerable time looking at illustrated material, they cannot identify relationships between information found in a data table and a graph of the same data. They certainly cannot interpret the graph to determine meaning, identify patterns, or establish the intended outcome of the graphical information.

Information graphics are pictures or illustrations that may or may not have embedded information hidden inside the illustration. In order for the reader to unlock the meaning of the illustration, they have to process, interpret, and interact with the hidden information. "Students need to become proficient in interpreting information graphics (e.g. graphs, tables, maps) because such graphics are used to manage, communicate, and analyse information." (Harris, 1996) This frustrating problem, which has plagued middle school teachers at my school, has led me to consider why my students do not have the abilities to interpret and find meaning in some information graphics.

REVIEW OF CURRENT LITERATURE

Even with the current climate for teachers, including high stakes testing, teachers often fall back into a traditional, lecture dependent style of instruction, despite research that children do not learn science concepts by reading a text, or by memorizing vocabulary (Olson & Horsley, 2000). Additional research has shown that certain types of instruction are more effective than other types at producing student gains (Prince, 2004; Lord & Orkwiszewski, 2006; McDermott & Redish, 1999). These interactive engagement methods, as identified by Hake, are examples of these more effective methods. These more effective methods involve getting the student engaged. Despite this research, most teachers continue to lecture primarily, even if they think they are using a more student-centered approach.

An interesting study in 2013 looked at the frequency of inquiry based science instruction occurring in high school classrooms (Capps & Crawford, 2013). The study evaluated twenty-six well-qualified and highly motivated fifth-ninth grade teachers from all over the country to determine the extent to which their instructional practices and views were reinforcing the nature of science. The goal of the study was to determine how often inquiry based instruction was occurring in science classrooms across the country.

Despite overwhelming previous research and widespread acceptance of inquiry based science teaching, researchers report that most science instruction in K-12 classrooms is teacher centered and dominated by teacher directed instruction (Alouf & Bentley, 2003). Would this new study find the same results of their research predecessors with the renewed push of STEM in the classroom? Researchers used many metrics to assess if inquiry instruction was occurring in teachers' classrooms. The methods of assessment included analysis of lesson descriptions,

classroom observations, videotape data, questionnaires, and interviews. The authors also sought to collect and analyze the teachers' statements about their own teaching with regard to inquiry and their understanding of the nature of science. The twenty-six teachers chosen from the application pool were based on their educational background, years of experience, college science coursework, presence of scientific research experience, and a willing administration. The teachers chosen for the study would be classified as highly qualified by many current educational standards. Teachers selected had an average of 11 years of experience and had on average 12 college level science courses. All but three of the participants had master's degrees, though two of the three were working towards their masters at the time of study. The participants were the best of the applicant pool. Surely, all of these highly qualified teachers were using inquiry based instructional methods in their classrooms.

Overall, the study found that only a few of the twenty-six teachers demonstrated a robust ability to teach "science as inquiry" based on the evidence that was collected. The authors concluded that, even though reform documents in the United States highlight the importance of inquiry and nature of science, relatively few teachers from their group of highly motivated teachers were actually teaching science as inquiry or about the nature of science. The research is not surprising; it is simply more data telling the same story that teachers are not teaching using research based strategies.

However, the research in this study is lacking in concrete quantitative results. The author makes large leaps in assuming that the group is composed of top tier science educators characterized as highly motivated based on external factors. The author would like you to assume that the group is a reflection of the best science teachers as a whole. In addition, the

various methods of data collection introduce the possibility for errors in the data that were not addressed in the literature. The story is clear and compelling however; teachers who believe that they are using inquiry-based practices are often not meeting an inquiry based instructional standard. One of the goals of my research was to evaluate my own teaching style to identify if my teaching would be characterized as inquiry based instruction.

In recent years, science education has placed increasing importance on learners' mastery of scientific reasoning (Duschl, Schweingruber, & Shouse, 2007). There exists an ever-increasing catalog of action research on how students are affected by different learning styles and how these students should be best tested to determine if they did or did not learn the material that was delivered to them. There are countless math, science, education, and psychology publications on how students see graphs, how they interpret pictures, how age effects understanding, how gender effects understanding. While there is a broad spectrum of research in science classrooms, little attention has been given to Earth Science concepts. Many studies look at the cognitive abilities of middle school students, but there is a scarcity of research in education in middle school sciences. The few research articles that do exist lack developed assessment tools and instructional strategies to use in the classroom.

As school districts and statewide testing requirements continue to place increasing importance into trying to gauge exactly what students know or more importantly what students have learned, the need for reliable, research validated assessment tools in the sciences will also continue to increase in importance:

> The growing emphasis on learner mastery presents a challenge for both developers, and users of assessments. Students and teachers are familiar with tests that offer assessment on content knowledge or a laboratory component that assesses the practical skills of a student. But assessing students' reasoning has required the development and

> validation of new kinds of assessments, supported by deep thinking about what constitutes scientific reasoning, which type of reasoning to assess, and what performances would indicate proficiency in a given type of reasoning (Rivet & Kastens, 2012).

Addressing this need for reliable testing instruments is occurring in some sciences much more quickly than in others.

In the early 1980's a group of physics researchers at Arizona State University were developing a test that measured the differences between Newtonian concepts and students' personal beliefs about the physical world. They gave this diagnostic test to numerous classes at the beginning of an introductory college physics course and again at the end of the course. The test was given in different teachers' classrooms, some who were more qualified than other teachers in physics instruction. The results showed the worst-case scenario for any teacher: students were not learning the physics concepts that had been specifically addressed during that same semester. The results further showed that the more qualified teachers did not have a noticeable effect on student learning (Halloun & Hestenes, 1985). This stark contradiction to what physics teachers believed was happening in their classes led many science education researchers to develop a more intuitive way to teach science, one that sought to emulate scientific practice by bringing those practices into the classroom. The modeling approach, developed by Wells and others, organizes the course content around a small number of basic models which describe basic patterns that appear ubiquitously in physical phenomena (Wells & al., 1994). So what began as a need for an accurate diagnostic tool in the classroom led to a teaching approach that has seen significant gains in college and high school classrooms (Hake, 1998; Laws & al., 1999). The reality is that, in large part, the diagnostic tools do not exist for

middle school teachers to accurately identify what cognitive and scientific abilities students have in their science classrooms.

Another goal of my study was to look at how an inquiry based instructional method might affect a student's ability to read and interpret information graphics. In my classroom, I have observed students struggle with graphs when they attempt them. Many students see a graphical question and they skip it altogether. I have noticed, in my classroom, that students who can understand graphs seem to do much better in the overall instructional setting of the science class. I have also noticed that students who struggle with graphs struggle with my science class. They also seem to struggle in many of their other classes. Could a student's graphical skill alone be an indicator of success in the classroom setting?

Research conducted in the 1960's concluded that graphicacy should be the fourth ace in the deck along with reading, writing, and arithmetic (Balchin & Coleman, 1965). What does it mean to be graphically literate or graphicate? Should education focus more on graphs? Students already look at content online that is embedded in pictures or diagrams. The use of multiple representations can help students develop different ideas and processes, constrain meanings and promote deeper understanding (Ainsworth, Bibby, Wood, 1997). If students are asked to organize their content from the classroom into multiple representations (graphically, mathematically, diagrammatically) the literature suggests that students may benefit in a deeper understanding of that content.

Often the struggle with graphical interpretations of information stem from students not being able to pull together information from multiple sources. Research shows that students encounter difficulty in integrating information from different sources (Case & Okamoto, 1996;

Demetriou & al., 2002). The idea that students have difficulty interpreting information found in graphs is not a new one. Research has shown that even late primary students have difficulty interpreting a number line. In fourth grade on the National Assessment of Education Progress(NAEP), students did no better than chance (1:4) on using a number line scale (Lowrie & Diezmann 2007) It is worth noting, however, that many studies indicate that graphical interpretation skill increases with age regardless of interaction with graphics (Wainer, 1980) (Liben & Downs, 1993) (Willett & al., 1988).

Researchers are working to determine what role information graphics play in understanding content and context in mathematical proficiency. Research has already shown that a relationship exists between information graphics and encoded math information. The same research also identified that the type of graphic had an effect on performance (Baker & al., 2001; Elia & al., 2007)). Previous work has shown that information graphics can be used to convey information both quantitatively, ordinal, and nominally through a range of perceptual elements (Mackinlay, 1999).

There is a wide range of elements employed to give meaning to information graphics. These elements include position, length, angle, slope, area, volume, density, color saturation, color hue, texture, connection, containment, and shape, among others (Cleveland & McGill, 1984). Information graphics can be broadly classified into six graphical languages which have unique spatial structures based on their perceptual elements and the encoding techniques that represent information (Mackinlay, 1999). Table 1 below introduces the six graphical languages into this research and offers a brief description of what each language asks from the individual.

A diagnostic tool to determine the level of students' understanding of embedded information found inside of graphical presentations is currently being developed. Original field research was conducted in 2007 on various test items for a graphic information diagnostic tool. Initial trialing was done with primary-aged students (2007, N = 172) (2008, N=796) in order for researchers to select items that met five characteristics. The researchers wanted items that : (a) required substantial levels of graphical interpretation, (b) required minimal mathematics content knowledge, (c) had low linguistic demand, (d) conformed to reliability and validity measures, and (e) varied in complexity (Diezmann & Lowrie 2008). Items were eliminated from the assessment tool after mass testing and interview testing to remove items that did not fit the desired characteristics. The test was further trialed over a three-year period in mass

Table 1. List of graphical languages with descriptions and examples

Graphical Language	Description	Examples
Map	Encoded through spatial location of the marks	Road map, topographic map
Axis	Encoded by placement of a mark on an axis	Number line
Retinal List	Encoded information not dependent on position	Graphics that feature color, shape, size, saturation, texture, orientation
Opposed Position	Encoded by a marked set positioned between two axes	Line chart, bar chart, plot chart
Miscellaneous	Encoded using variety of graphical techniques	Pie chart, Venn diagram
Connection	Encoded by node objects with link objects	Network, dichotomous key

testing situations and in interview situations to ensure content reliability and validity. The resulting test, the Graphical Languages in Mathematics, was shown to be useful in identifying student weaknesses in interpreting information graphics. The research suggests that mathematical proficiency is fundamental to students' abilities to interpret information graphics

and that just as the mathematical and linguistic demands of test items impact on performance, so too does the graphical component (Diezmann & Lowrie, 2008).

As I read the published articles on the GLIM I was instantly pulled in by its potential impact on students in my classes. I wanted to see how the researchers had used it in the field to affect graphical literacy. I was unable to find any research indicating the tool was used to measure what influenced graphical reasoning. Since I could not find the research that I was looking for I decided that I should use the GLIM to look for changes in graphical proficiency during the year.

MATERIALS AND METHODS

The purpose of this study was to determine the effect of a student centered learning environment on middle school students in an earth science classroom to see if this environment improved student ability in graphing and scientific reasoning. The data used in this study was collected during one academic year at the same middle school in southeast Baton Rouge. A breakdown of the middle school providing population totals, ethnic background, and students receiving free or reduced lunch (an indicator of poverty) is in Table 2.

Table 2. Demographic make-up of school population (eSCHOOL PLUS)

Demographic and Lunch Status of Woodlawn Middle School, 2014 Total School Population = 1014			
Ethnicity	African American = 642 (63%)	Caucasian = 262 (26%)	Other = 110 (11%)
Lunch Status	Free/Reduced = 808 (80%)	Full Price = 206 (20%)	
Gender	Female = 487 (48%)	Male = 527 (52%)	

The specific data in this study comes from six different classes of eighth grade students in one teacher's classroom (N=114). Four of the classes involved in the study were characterized as "regular" classes. Students in regular classes are part of the general population of eighth grade students, meaning they have not been identified prior to eighth grade as needing considerably more support than regular classes offer (special education) nor have they been identified as having an above average ability or aptitude in previous years of school or on standardized testing in any subject matter. There are students in these classes that have been identified as needing special accommodations in the classroom, which may

require additional testing time or special conditions for class instruction or testing. There were a total of eleven of these students in all of my classes. Eight of the students were in the experimental group and three of the students were in the control group, but the students were not given any special considerations during the collection of this research data.

Students who are identified as having an above average ability in class material or on standardized tests are often given the option to participate in classes that aim to provide a more rigorous workload by either going through material in more detail or by going over material at a faster pace than the regular classes that are offered. Woodlawn Middle has two programs for students who qualify for an accelerated classroom setting: gifted and Great Scholars. Gifted students, who I am only identifying for purposes of distinguishing Woodlawn's total population, must qualify for classes based on third party aptitude and psychological testing. None of the students in this study are part of the gifted program at Woodlawn.

There is also a Great Scholars program. Students enrolled into the Great Scholars classes qualify based on above average performance on either end of course standardized tests required by the state (LEAP, iLEAP) or by consistent above average performance on end of unit tests in the classroom. Students are then given an aptitude test by the school to qualify them for the Great Scholars program. Two of the classes in this study were characterized as Great Scholars (N=41): one group was in the control group and one group was in the experimental group. The entire sample was one hundred and fourteen students and was broken down by block, gender, ethnic background, and regular/Great Scholars groups as shown in Table 3.

Table 3. Demographic make-up of study population. (eSCHOOL PLUS)

Group	Total	Males	Females	African American	Caucasian	Other	Regular	Great Scholars
Control	56	29	27	36	15	5	34	22
Experimental	58	32	26	34	18	6	39	19

The school uses a block schedule where students have eight different classes over a two-day period. The days are broken down as A-day students and B-day students. For this study, students on A-day were in the control group and received traditional instruction. There were three periods of traditional instruction taught on A-day with fifty-six students in those combined classes. Students on B-day were in the experimental group and received a mixture of different instructional methods. There were three periods of student centered learning instructional methods and fifty-eight students participated in the experimental group. Both groups received the same amount of class time for instruction during a two-week period. Each group participated in five science classes during the two-week period and each class lasted ninety minutes. The control group was taught using traditional instructional methods (lecture, PowerPoint, bookwork) and the experimental group was taught using mixed methods of student centered instructional methods. The methods used in the experimental group were heavily influenced by modeling techniques that were learned at a two-week modeling workshop.

Before teachers can effectively utilize modeling in the classroom, they must attend a teachers' modeling workshop where they learn the modeling cycle as they work through major portions of the curriculum as both a teacher and a student. If I was going to implement a

modeling instructional strategy in my classroom, I would have to attend a training to prepare for this study. Two modeling courses were available locally prior to my research project and neither of those courses was in earth science. In fact, there was not a course available anywhere in the United States that would allow me to learn the modeling framework for earth science, for the simple reason that one does not exist. Therefore, with the hope that I would be able to marry physical science modeling in my earth science classroom, I attended a two-week physical science modeling workshop in New Orleans.

During the two-week physical science modeling course I was able to find some overlap with my earth science curriculum. The first and second instructional units in the East Baton Rouge Parish curriculum for earth science are a review of the scientific method and scientific measurement. In the modeling framework, the scientific method is reinforced through the experimental design of the activities. Examples of the instructional materials used from modeling can be found in appendix A, appendix B, and appendix C. Most of the modeling activities began with “what can you observe about this item or these items in a system?” After students have made observations and shared those with their group, students are asked, “What can you measure about this object or system?” This method seemed to reinforce experimental design and there is a significant portion of the physical science modeling that has the students identifying the structure of measurement. After students have developed models for measurement, they discover the idea of unit analysis, significant digits and uncertainty in measurement. They also learn how to represent these relationships mathematically, diagrammatically, and graphically. The modeling curriculum continues with measurement. Unfortunately, the earth science curriculum and the physical science curriculum did not have

any other similarities. I was confident that after the modeling training I would be able to convert some of my earth science curriculum into models that resembled the modeling framework.

To assess my students learning, I used the Middle School Graphing Inventory (MSGI) and the Lawson's Classroom Test of Scientific Reasoning (CTSR). The MSGI is a twenty-seven item test that assesses a student's ability to read and interpret information graphics. Most of the items contained in the MSGI were used with permission from the creators of the Graphical Languages in Mathematics (GLIM) test, which was discussed earlier in the section. The MSGI measures a student's ability to interpret six different types of information graphics: maps, single axis items, retinal list items, miscellaneous items, and connection items. Table 1 above provides examples and descriptions of each of the graphical languages.

Lawson's Classroom Test of Scientific Reasoning (CTSR) was used to identify students' scientific reasoning abilities. The CTSR, developed by Anton Lawson, has been validated and used extensively as a research instrument in measuring middle school, high school, and college students' scientific reasoning abilities in different science disciplines (Lawson, 1978; Lawson & al., 2000; 2007; Moore & Rubbo, 2012). The test is a useful diagnostic tool to identify a student's level of science reasoning; specifically a student's formal operational ability. Lawson defines formal operations as:

> those reasoning processes that guide the search for and evaluation of evidence to support or reject hypothetical causal propositions. These operations are used in the isolation and control of variables, the combinatorial analysis of possible causal factors (combinatorial reasoning), the weighing of confirming and disconfirming cases (correlation reasoning), the recognition of the probabilistic nature of phenomena (probabilistic reasoning), and the eventual establishment of functional relationships between variables (proportional reasoning). (Lawson, 1978).

Different research studies have chosen to interpret the results in various ways, but most are similar in that they distinguish between low scientific or concrete reasoning ability, transitional reasoning ability, and formal operational reasoning ability. In my classroom, the CTSR was used to identify initial science reasoning abilities of groups and it was used to identify changes in science reasoning ability during the school year.

These two tests were used as the assessment instruments for my study. The CTSR was administered in the last week of August as a pretest and the MSGI was administered the following week. The treatment began at the beginning of September and concluded in the middle of November before students had a fall break. The treatment stopped primarily because the IE strategy used took longer to deliver than the traditional treatment, so I was getting behind in the classroom pacing for earth science. Furthermore, earth science materials do not exist in the modeling framework, and I did not have time to create new IE materials in Earth Science for the rest of the year.

With the initial treatment concluding, I administered the two tests again as midterm tests at the beginning of the second semester. At the time, I was hopeful that I would be able to implement additional IE strategies during the second semester, but I was unable to do so due to the amount of time that would be required to develop IE materials for the second semester. I gave the two tests again in the beginning of May, at the end of the school year, to look at the long-term retention of the material from the study.

For the purpose of this study, I chose to use one day of instruction as my control group (A day) and the other day of instruction, (B day) would serve as my experimental group. The control group would receive traditional classroom instruction. I felt confident in my ability to

deliver traditional instruction to the control groups because this was my fifth year teaching the same course. Students in this class went over overhead computer slides about the scientific method and measurement during the treatment period. Students were also given lab assignments that had them practice measurements of length, volume, and mass. In addition, students in the control group practiced reading measurement tools with handouts and an online virtual scale. Students were given a measurement test at the conclusion of the instruction. In addition, students were given the midterm graphing test and the midterm science reasoning test (CTSR) after they returned from winter break.

Students that participated in the experimental group received an initial PowerPoint presentation on the scientific method and an introduction to scientific measurement. While the control groups were completing lab activities during their instructional block, the experimental groups were divided into three person teams and completed activities and assignments to the modeling framework from physical science. Unit 1, Chapter 1(U1.1) “Measuring the Measuring Tool” was the experimental groups’ first introduction to interactive engagement. Students were asked during the activity to observe four common items: an ink pen, a pencil, a paper clip, and a standard eraser. Students were asked to make comparisons using the items as measuring tools. Students were asked to measure each of the items using the other items as the measuring device. Students were also asked to measure some common items in the class. During the activity students struggled with appropriate unit conversions and the students also had to make decisions about how to correctly measure with their items. The instructions were purposefully ambiguous to introduce those conversations into the groups. Students asked the teacher for input about the right way to do the measurement, but the

activity is meant to identify that, like experimental design, there are numerous scenarios that can be correctly used to complete the activity. Students were then asked to describe their experimental procedures in the accompanying handout.

At the completion of the measurements, which took most of one class period, the students were asked to transfer their results to a large whiteboard and include all of their measurement comparisons. After most groups were done copying their data to the whiteboard a “board meeting” was conducted. In the board meeting students were asked again to make observations. They were asked to look at the data tables from the other groups and identify characteristics of the information of other groups as well as their own. Students were prompted to identify at least one similarity between all of the whiteboards and to identify one thing that was different between groups. Students at this point were also encouraged to think about questions they might ask other groups about the information on their boards.

Another important component of the modeling cycle occurs when groups have to present their data to their peers. Many of the tasks in the activity ask students to do things that are not common to most of their previous classrooms. While students are occasionally asked to record data in other classes, they are not typically asked to orally present the data that has been collected. Students are regularly asked to identify similar items in a text or a paragraph, but are not typically asked to pull similarities and differences from multiple sources. It is recognized in educational literature that when students are asked to identify similarities and differences that higher order thinking is required of the student and this type of activity is encouraged in classrooms (Marzano & al., 2001). Students did not consistently represent their data from table to table on the whiteboards, so students had difficulty figuring out how to read

each table. Students made mistakes in measurement and in collection, but many of those problems were identified in the board meeting *by the students.* Many of the students struggled with the activity because they were not used to activities without instructions and procedures being explicitly spelled out. During the board meeting students are asked to determine as a group if there was a best way or most consistent way that procedures were done. In addition students are asked to consider what problems they can identify with the design of the activity. During the board meeting the groups decided that having a similar consistent procedure for measurement was necessary and this became the group's model. Once all of the groups had presented and identified observations about other data tables they completed an assignment that reinforced the ideas of consistency in measurement, but also introduces a new concept in converting a measurement using unit analysis.

This basic structure was followed with subsequent activities in the physical science modeling curriculum. Students went through the second activity, precision in measurement. Homework assignments were given at the conclusion of the second activity and the groups were all introduced to the first activity in the second unit, wingspan versus height. In this activity students are introduced to plotting points onto a graph and how to find a line of best fit. The Great Scholars experimental group completed the activity, while the regular groups were able to begin the activity. The discussion following wingspan versus height was not introduced, nor were the results individually white boarded. Due to curriculum conflicts between the physical science modeling and the required content on the district's earth science pacing guide, I had to move on to the following units of instruction in the earth science curriculum. Later in the month, the experimental groups were able to whiteboard and have a

board meeting identifying similarities and differences between mineral concept maps that each group developed. The experimental groups' treatment concluded in the third week of November when the students left for Thanksgiving break.

During the school year I had some students who entered my classroom after the school year began. I also had students who transferred to other schools during the experiment, or did not take both pretest and posttest evaluations. Those students were omitted from the study. Ten students in the sample participated in the pretest and posttest evaluations, but did not take the midterm evaluations. Those students' pretest and posttest scores were used for this study, but were omitted from analysis of pretest/midterm/posttest evaluations.

DATA AND RESULTS

Statistical analysis of pretests, midterms, and posttests collected during this study was conducted using mean classroom scores. An analysis of variance test (ANOVA) and two tailed t-tests (independent and paired) were used to compare and contrast groups. Mean classroom scores were used to determine overall performance of participants in a given population. If possible differences were identified in student scores from different populations an independent t-test or an ANOVA was performed. For the statistical analysis in this study p-values of 0.05 or lower were the acceptable thresholds for showing groups or scores were dissimilar.

Student raw gains were also analyzed for statistical variance either within groups or between groups. Raw gains identify the change in a specific assessment each time the assessment is given. Raw gains are determined by subtracting the pretest score from the posttest score (posttest-pretest). Raw scores individually show if a student's performance improved or declined from the previous administration of the assessment. In aggregate, raw gains can be used to identify trends where groups of students in a population improved or declined from the previous administration of the same test.

Students participating in the study were given the Lawson's Classroom Test of Scientific Reasoning (CTSR) and the Middle School Graphing Inventory (MSGI). The tests were administered at the beginning of the 2013-2014 school year. All six classes were given both tests. I teach two different groups of students at my school, Great Scholars (GS) and regular students (Reg). Great Scholars and regular students are all in the eighth-grade, with the

exception that GS classes move at a faster pacing or learn concepts to a greater "depth of understanding."

Great Scholar and Regular Groups Pretest Scores on MSGI and CTSR

Figure 1 below shows the average pretest scores for groups of students on the MSGI. Scores between groups were analyzed at the beginning of the study to confirm that the groups being compared were similar in composition.

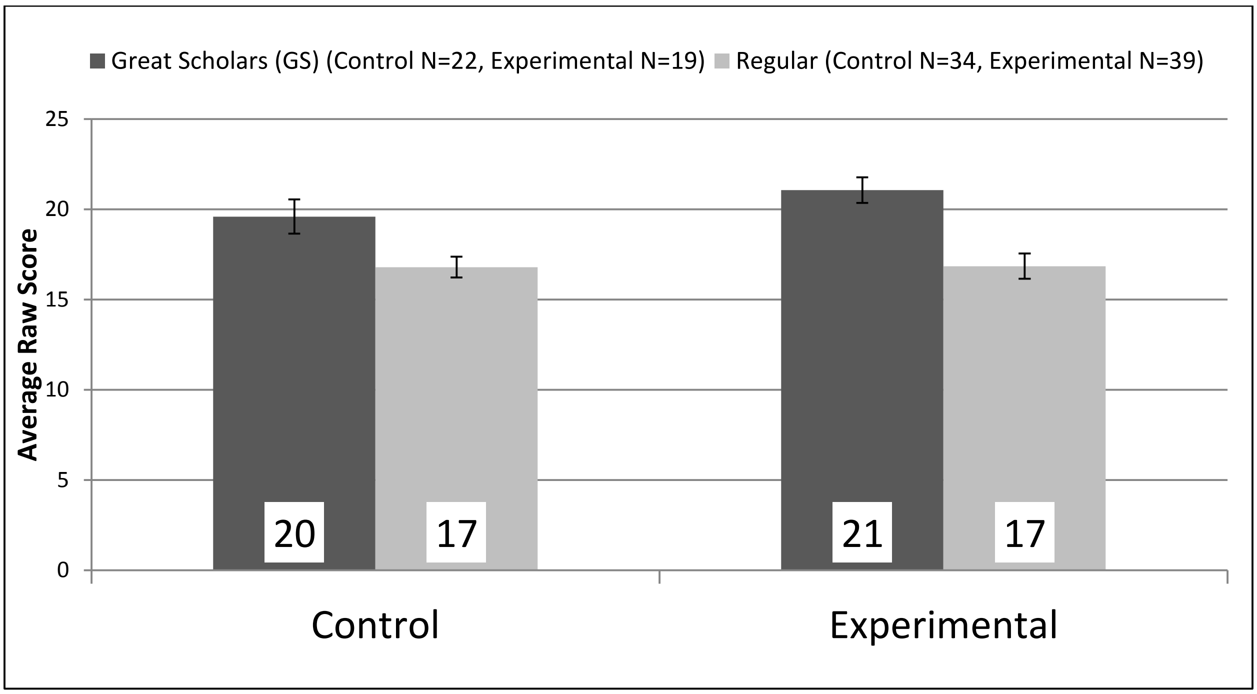


Figure 1. Great Scholars (dark gray columns) and Regular (light gray columns) students average raw pretest scores on MSGI. Control group at left (N = 56) and experimental group (N = 58) at right.

For the Great Scholars, the control group had an average score of (19.6 ± 0.9) and the experimental group had a pretest average of (21.1 ± 0.7). The regular students had lower average scores; the control group at (16.8 ± 0.6) and the experimental group at (16.8 ± 0.7). An ANOVA on the four groups shows them to be significantly different ($p < 0.001$). The two classes of Great Scholars students were compared to one another using a two-tailed t-test ($p = 0.2$).

The four remaining groups of regular students were analyzed using an ANOVA ($p = 0.95$). So, while the Great Scholars and regular students have different graphing abilities, the subgroups of Great Scholars (Great Scholars control & Great Scholars experimental) and regular students (regular control & regular experimental) are not statistically different and have similar graphing abilities.

Lawson's Classroom test of Scientific Reasoning was also given to all students at the beginning of the school year to identify students' scientific reasoning abilities. All classes average raw score on the test were calculated with the GS control group (9.5 ± 0.8) and the GS experimental group (8.9 ± 1.0) having average scores statistically higher ($p < 0.001$) than the regular control group (6.9 ± 0.4) and the regular experimental group (5.9 ± 0.4). Figure 2 below shows a bar graph comparing the average scores on the pretest of each of the groups. So we

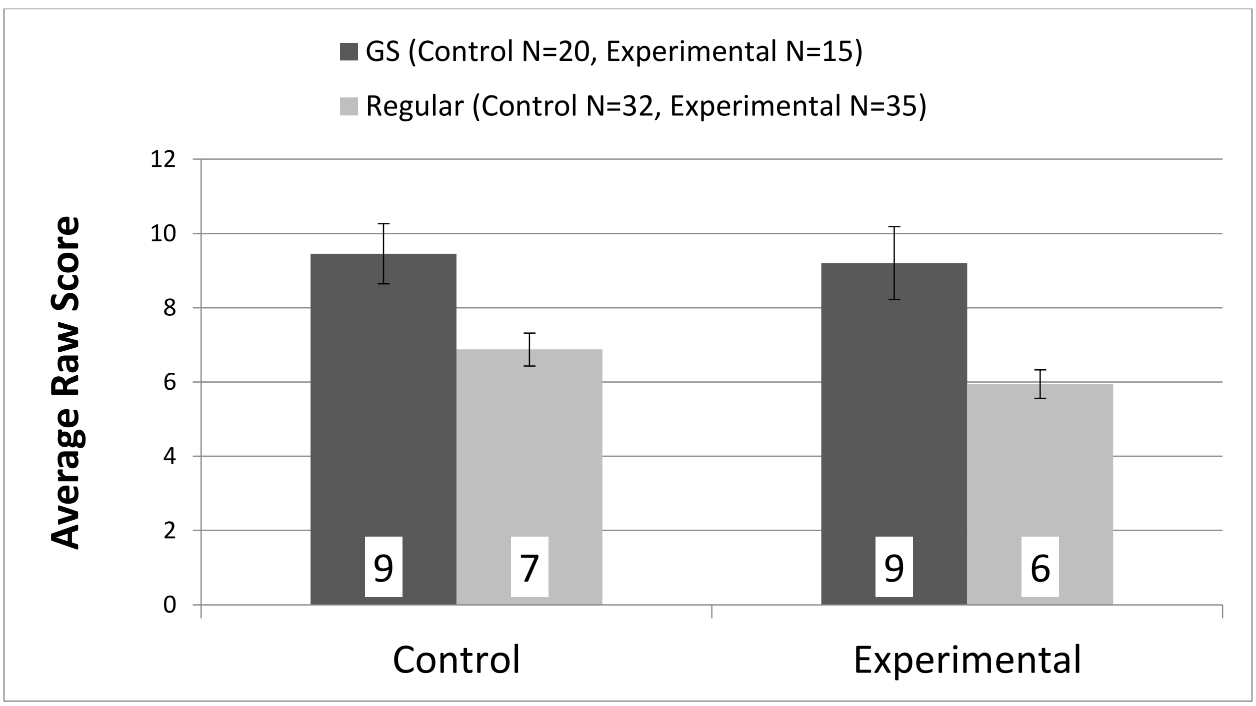


Figure 2. Great Scholars (dark gray columns) and Regular (light gray columns) students average raw pretest scores on CTSR. Control group at left (N = 52) and experimental group (N = 50) at right.

see that pretest performance on the CTSR, and the groups' science reasoning abilities, are very different between the GS group and the regular group. The two Great Scholar classes were compared and were not statistically different ($p = 0.8$), nor were the four regular groups different ($p = 0.1$) when compared using an ANOVA.

We see that the Great Scholar groups do not have the same graphing ability or science reasoning ability as the regular groups and cannot be compared as similar populations. We do find that within the GS group that the population of students is similar and we also find that within the four regular groups we do not see statistical differences and we can compare those as similar groups.

Great Scholar and Regular Groups Results on MSGI

In order for us to evaluate if teaching style impacted students we will look at the average scores of each group over time. In Figure 3 below, we see how the Great Scholar groups' graphing skills are affected by the type of instruction they receive. The GS control group on the left side of the graph received traditional instruction and the GS experimental group on the right side of the graph received different types of interactive engagement that were pulled from modeling physical science during the first semester. Mean scores for the GS experimental group move from a pretest mean score of (21.1 ± 0.7), to a midterm mean score of (21.4 ± 0.5), to a posttest mean score of (22.6 ± 0.5), but a statistical analysis (ANOVA) reveals that differences were not significant ($p = 0.15$). The control group, which received traditional instruction during the first and second semester, saw significant gains ($p = 0.03$) in average scores from pretest (19.6 ± 0.9), to midterm (19.6 ± 0.8), to posttest (22.1 ± 0.4). This is interesting because the control group sees significant gain from midterm to posttest, but it

does not happen during the course of the treatment in my classroom. This will be discussed later in the results.

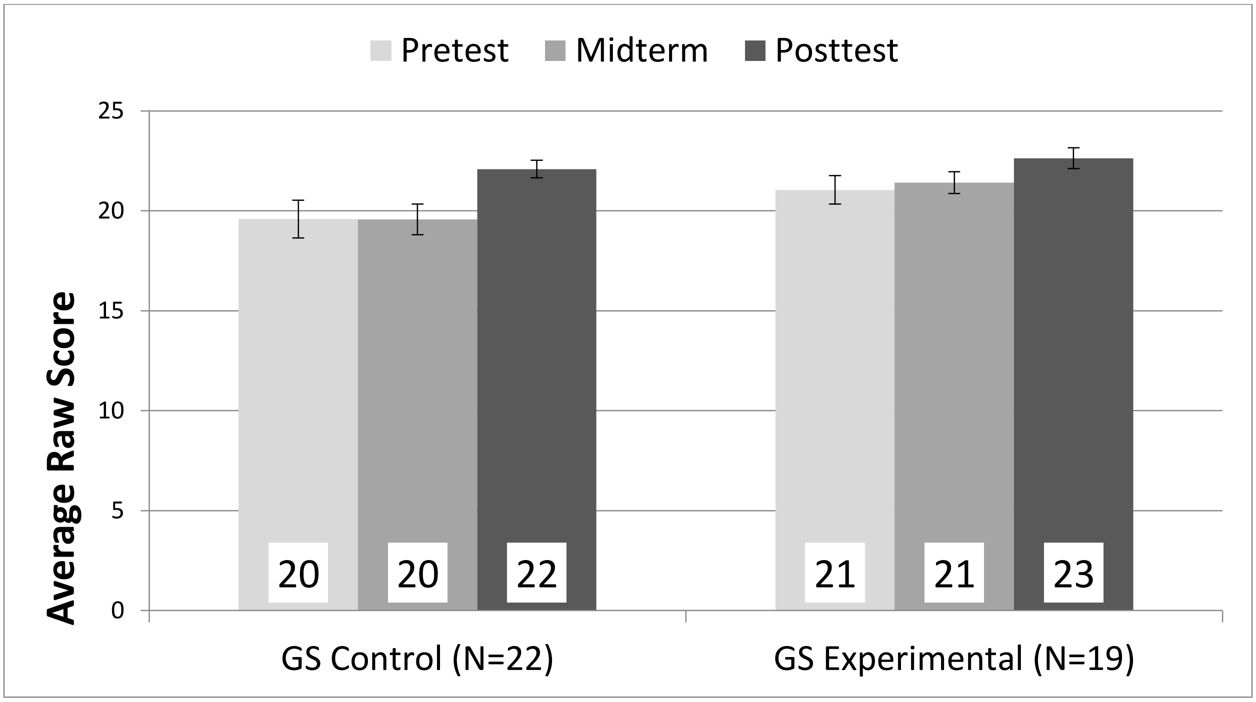


Figure 3. GS control (left) & experimental (right) groups average raw scores on pretest, midterm, and posttest over time on MSGI.

In the previous figure we saw that the GS experimental group did not see significant gain on the graphing inventory from mean pretest scores to mean posttest scores. Figure 4 below shows how regular students' graphing ability changed as a result of the type of instruction they received.

The regular control group that received traditional instruction during the first semester had an average pretest score of (16.8 ± 0.6) and an average midterm score of (17.6 ± 0.7). A paired t-test of the two groups did not reveal a significant difference at the end of the semester ($p = 0.12$). End of the year average posttest results again see a slight increase in the mean score

(18.7 ± 0.7) and a two tailed paired t-test between average pretest scores and average posttest scores ($p < 0.01$) shows us that the average scores are statistically different.

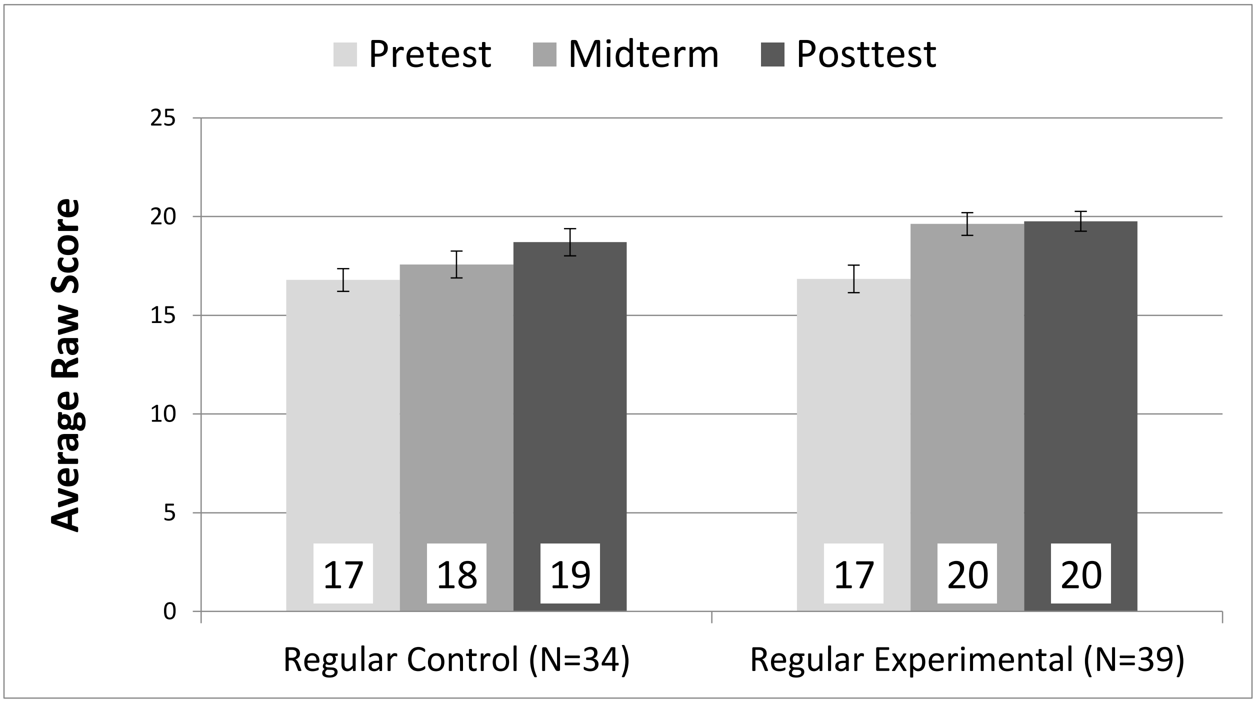


Figure 4. Regular control (left) & experimental (right) average raw scores on pretest, midterm, and posttest over time on MSGI.

The regular experimental group, which received interactive engagement instruction only in the first semester, had an average pretest score of (16.8 ± 0.7) and an average midterm score of (19.6 ± 0.6). A two tailed paired t-test of the two samples ($p < 0.001$) identifies that the average score of the regular experimental group improves significantly. The regular experimental group's average posttest score (19.8 ± 0.5) is also statistically significant growth ($p < 0.001$) when compared against average pretest scores.

So we see that both of the regular groups had statistically significant gains, but the timing of the growth is important. The regular experimental group saw growth in the first semester using interactive engagement strategies in the classroom where the regular control

group did not have growth in the first semester. The regular control group, which received no treatment in the first semester, saw no gains in the first semester.

The regular control group behaved similarly to the Great Scholars control group in that both groups saw a statistically significant growth from pretest to posttest with no intervention. Both groups also showed no significant changes during the first semester and showed a statistically significant change during the second semester of school. We will discuss shortly these results by examining the math backgrounds of the students.

Great Scholar and Regular Groups Results on CTSR

In addition to using the graphing test to look for differences in graphing ability over time, the Lawson's Classroom Test of Science Reasoning (CTSR) was used to identify if students' ability to reason in science was effected by the instructional method in the classroom. Figure 5 below shows Great Scholar groups' average science reasoning scores over time on the CTSR.

The GS control group, which received traditional instruction during the treatment period had an average pretest score (9.5 ± 0.8) compared to an average midterm score (10.2 ± 0.8) and an average posttest score of (10.1 ± 0.9). Analysis of variance was run on the pretest, midterm, and posttest and was not found to be statistically significant ($p = 0.8$). The GS experimental group, which received interactive engagement lessons in the classroom during the first semester of school saw an average pretest score of (8.9 ± 1.0), an average midterm score of (10.4 ± 1.0), and an average posttest score of (9.9 ± 1.0). An ANOVA run on these scores showed the differences to not be significant ($p = 0.56$).

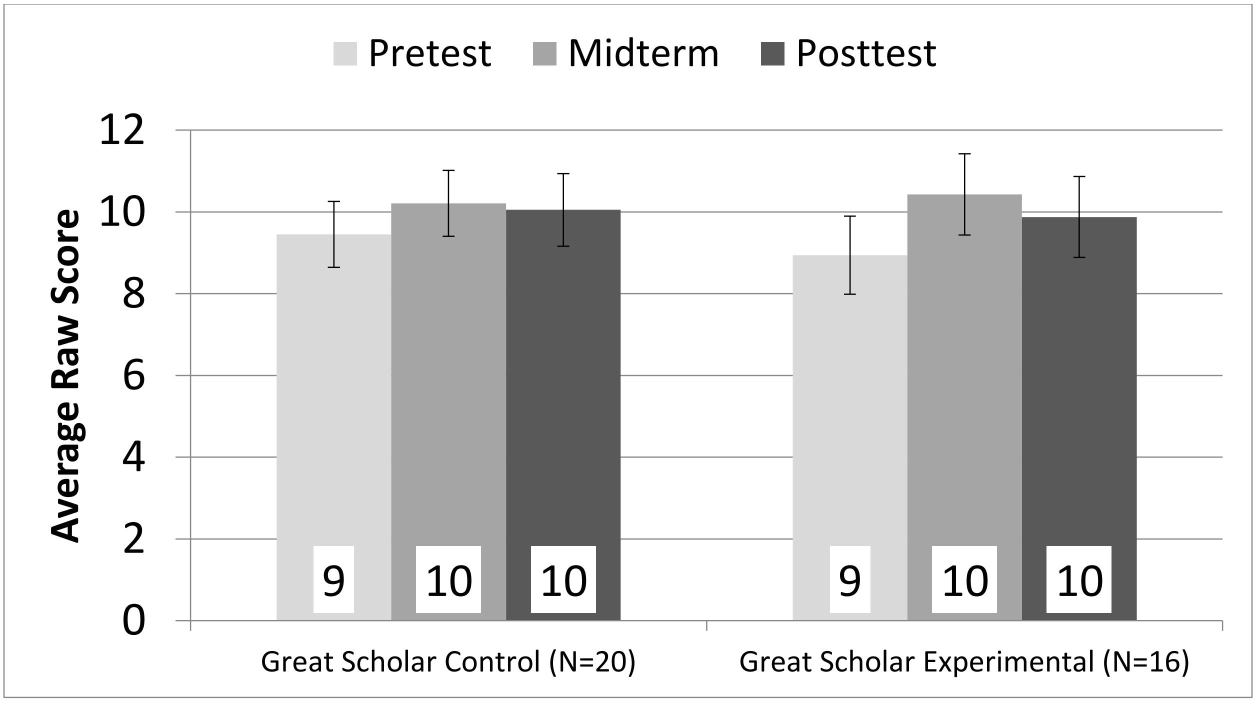


Figure 5. Great Scholar control (left) & experimental (right) groups average raw scores on pretest, midterm, and posttest over time on CTSR.

In the previous figure we did not see any large differences over time in GS science reasoning ability due to instructional method. Figure 6 below compares the regular groups' change in science reasoning ability (CTSR) over time due to instructional method. The regular control group had an average pretest score on the CTSR of (6.9 ± 0.4), an average midterm score of (6.7 ± 0.4) at the end of the first semester, and an end of the year posttest score of (6.6 ± 0.5). An ANOVA analysis of these results revealed no statistical difference in the students' scores ($p = 0.89$).

The regular experimental group had an average pretest score of (5.9 ± 0.4) and an average midterm score after treatment of (6.7 ± 0.5). The end of the year average posttest score was (7.6 ± 0.5). An ANOVA between the regular experimental group's test scores yielded

a p - value of (p = 0.03) indicating that there was statistically significant growth in the regular experimental groups' scientific reasoning ability over the year.

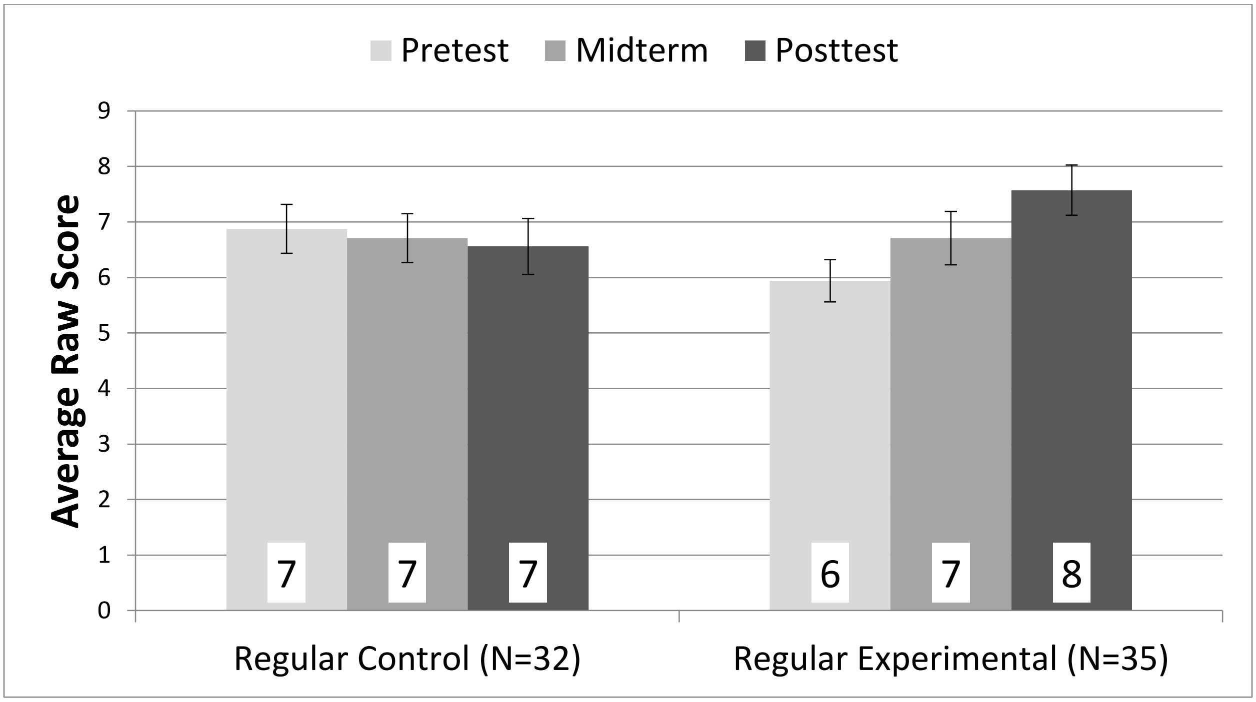


Figure 6. Regular control (left) & experimental groups average raw scores on pretest, midterm, and posttest over time on CTSR.

Using the graphing inventory (MSGI) and classroom test of science reasoning (CTSR) to evaluate the effectiveness of interactive engagements methods in the class produced unexpected mixed results in this study. Similar classroom testing has been done in physical science classrooms and in high school classes, but there is not significant research that exists about using these assessments in middle school to measure graphing ability and assess instructional method. Students in both of the regular groups saw significant gains in graphing ability over time. Additionally, the students in the regular group who were taught using interactive engagement strategies saw a significant gain in science reasoning ability over the course of the year.

The interesting and unexpected results occurred in the GS and regular groups that were taught using traditional instruction. Both of these groups saw significant gains in graphing ability, but they occurred in the second semester after the experiment was completed. Additionally, both groups only participated in traditional instruction and were not expected to show significant results. Is there anything else that could be affecting students' abilities to read and understand graphical content besides their science class? Earlier conclusions made by Diezmann & Lowrie in testing their GLIM math test suggested that mathematical proficiency played a critical role in students' ability to read and interpret information graphics.

Math Groups Pretest Scores on MSGI and CTSR

Figure 7 below identifies the math classes that my students were enrolled in at the time of the study as well as the math class that they had the prior year. Great Scholar groups took

	7th Grade Math Course	8th Grade Math Course
Great Scholars	High EOC =	GS Algebra (N = 19)
	Pre-Algebra (8th grade material) (Graphing -2nd Sem)	
	Low EOC =	GS Pre-Algebra(Repeat) (N = 22) (Graphing -2nd Sem)
Regular Students	**7th Grade Mathematics** (limited graphing)	Pre-Algebra (N=73) (Graphing -2nd Sem)

Figure 7. Student math courses in 7th & 8th grade. Identifies when students' math courses specifically teach graphing content

eighth grade math in their previous year of school. At the end of the GS students' seventh grade math class, the students take an end of course test (EOC) covering eighth grade pre-algebra to determine if students meet proficiency requirements. Students who meet requirements move on to algebra in their eighth grade year; students who do not meet requirements do not move on to algebra, but instead repeat eighth grade pre-algebra as an

elective that only meets every other day. All other students at Woodlawn Middle School get mathematics instruction daily. All of the students in regular eighth grade classes are enrolled in pre-algebra.

We will now look at the effect of math class on a student's graphing ability on the middle school graphing inventory. We will also look at the effect of math class on a student's ability on their scientific reasoning skills. Student groups will now be regrouped to reflect their current math classes. All students in regular pre-algebra make up the population of one group, the regular pre-algebra group (N = 73). All students in Great Scholars program who are repeating pre-algebra make up the population of the second group, GS pre-algebra (N = 22). The final group is made up of Great Scholar students who are enrolled in algebra (N = 19).

Figure 8 below shows the results of average pretest scores for each of the math groupings. The regular pre-algebra group had the lowest mean pretest score of (16.8 ± 0.5),

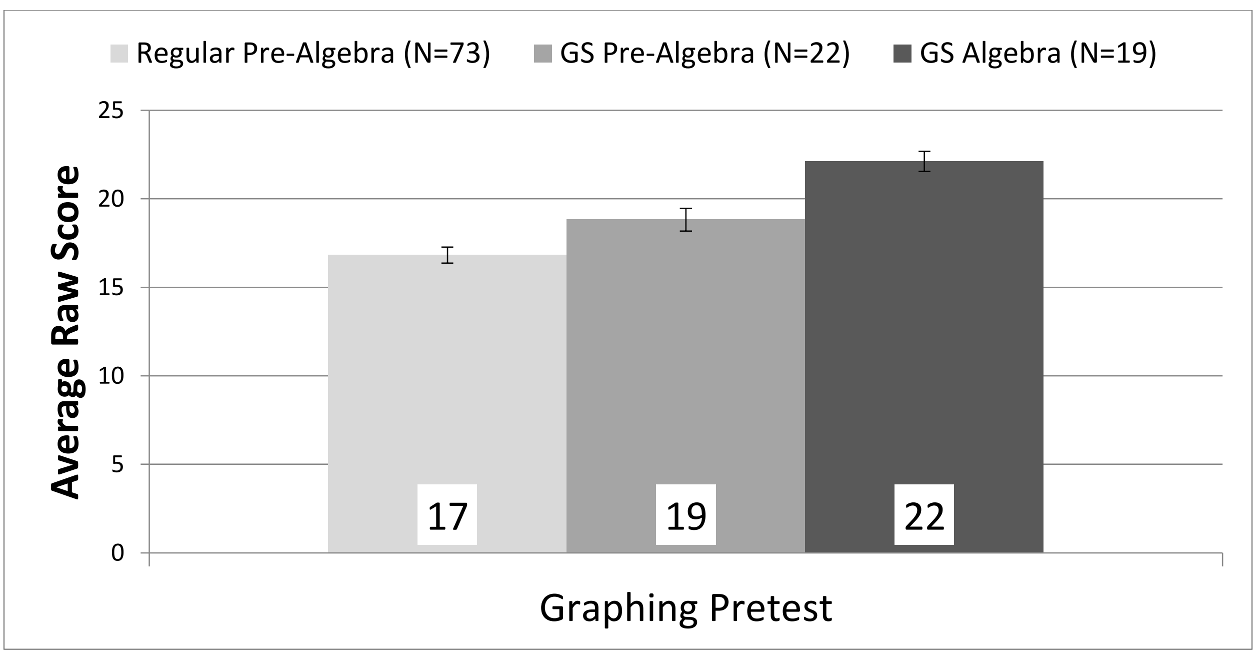


Figure 8. Average raw pretest scores on MSGI by math class. Regular Pre-Algebra (light gray, N = 73), GS Pre-Algebra (medium gray, N = 22), and GS Algebra (dark gray, N = 19).

followed by the GS pre-algebra with an average pretest score of (18.8 ± 0.9), and the GS algebra group had the highest pretest average on the MSGI with a score of (22.1 ± 0.6). An ANOVA analysis of these three groups shows their initial ability in graphing to be significantly different ($p < 0.001$).

Figure 9 shows the average pretest CTSR scores of each math group to measure their prior scientific reasoning. The average pretest score of the regular pre-algebra group is the lowest at (6.4 ± 0.3), followed by the average pretest score of the GS pre-algebra group at (8.2 ± 0.7), and the GS algebra pretest score at (10.6 ± 1.0). ANOVA analysis confirms these differences are significant ($p < 0.001$). The groups, separated by math class, are statistically different in both graphing ability and scientific reasoning with the strongest reasoning and graphing ability coming from the GS algebra groups. The GS pre-algebra group has a lower graphing and scientific reasoning ability and the lowest group, the regular pre-algebra students.

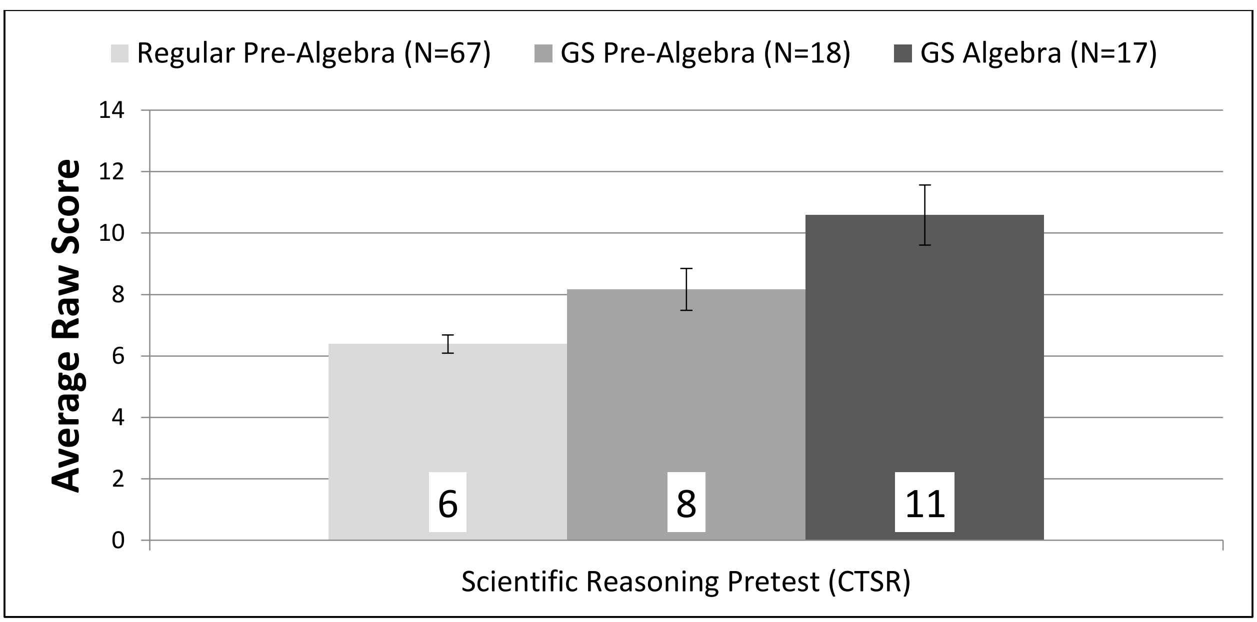


Figure 9. Average raw pretest on CTSR by math class. Regular Pre-Algebra (light gray, N = 67), GS Pre-Algebra (medium gray, N = 18), and GS Algebra (dark gray, N = 17).

Math Groups Results on MSGI

We want to know if students' math classes have an effect on the graphing ability and scientific reasoning of each group. Figure 10 below shows how math classes have affected each group's ability to graph. In the first column we see that the regular pre-algebra group has an

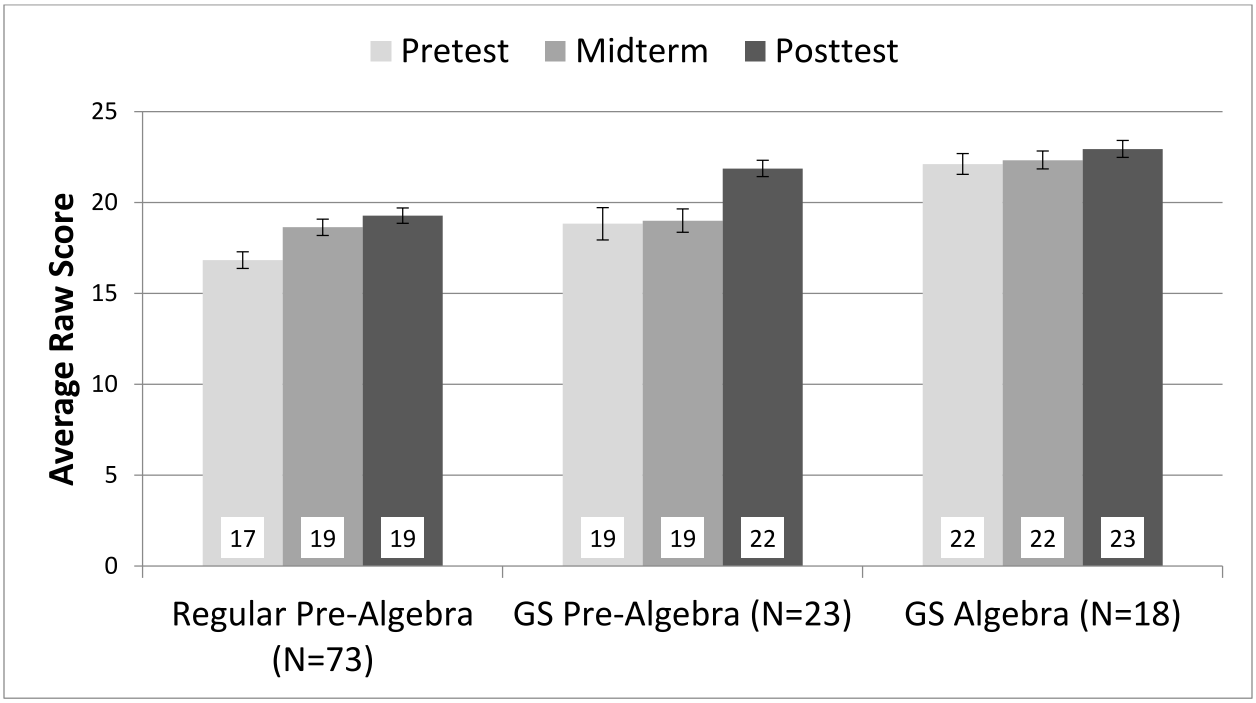


Figure 10. Math groups results over time on MSGI. Regular Pre-Algebra (left), Great Scholar Pre-Algebra (center), and Great Scholar Algebra (right) average raw scores on pretest (light gray), midterm (medium gray), and posttest (dark gray).

average pretest score of (16.8 ± 0.5), an average midterm score at (18.6 ± 0.5), and an average posttest score of (19.3 ± 0.4). An ANOVA on the three scores showed that the gains in the regular pre-algebra group were statistically significant ($p < 0.001$). The GS pre-algebra group that is repeating algebra had an average pretest score of (18.8 ± 0.9), an average midterm score of (19.0 ± 0.6), and an average posttest score of (21.9 ± 0.5). An ANOVA of the group's three scores confirmed the gains by the GS pre-algebra group was statistically significant ($p = 0.003$). The Great Scholars algebra group had an average pretest score of (22.1 ± 0.6), an average

midterm score of (22.3 ± 0.5), and an average posttest score of (22.9 ± 0.5). An analysis of variance identified that the GS algebra growth was not statistically significant ($p = 0.5$).

We saw statistically significant growth by both pre-algebra groups when looking at their ability to perform graphical tasks over time. Three things are interesting in the results of this data. In the GS pre-algebra group, we see no growth from pretest to midterm, when I was using interactive engagement in the classroom. What we see is all of the GS pre-algebra growth occurred during the second semester when I was not doing activities that required graphing skill. The second semester of math for the GS pre-algebra group focused on graphical tasks. This is an indication that the science class did not produce the gains we see in student learning, but the math class is what likely caused the growth on the MSGI.

We see that the regular pre-algebra groups saw statistically significant growth in the first semester only ($p = 0.006$). The gains seen during the second semester were analyzed using a two-tailed paired t-test and were not shown to be statistically significant ($p = 0.08$). The regular pre-algebra students had statistically significant gains during the first semester, when math classes had not introduced graphical skills heavily. The graphically focused work in the math classes did not occur until the second semester. This is a strong indication that the regular pre-algebra students showed growth due to the interventions done in the science classroom, and not in their math classes.

We also see that in each of the lower math groups (regular pre-algebra and GS pre-algebra) that they perform similarly to the pretest scores of the more advanced group by the end of the year. The GS pre-algebra students' gains catch them up to the pretest achievement

level of the algebra group by the end of the year. Similarly, we see that the regular pre-algebra groups' gains catch them up to the pretest performance level of the GS pre-algebra students.

Math Groups Results on CTSR

We used the CTSR to measure if math classes had an effect on the students' scientific reasoning abilities. Figure 11 shows the average pretest, midterm, and posttest results of each math group over time. The regular pre-algebra group had an average pretest score of (6.4 ± 0.3), an average midterm score of (6.7 ± 0.3), and an average posttest score of (7.1 ± 0.3). An ANOVA of the regular pre-algebra group shows us that the scores in this group are not statistically different over time ($p = 0.29$). The GS pre-algebra was analyzed next and

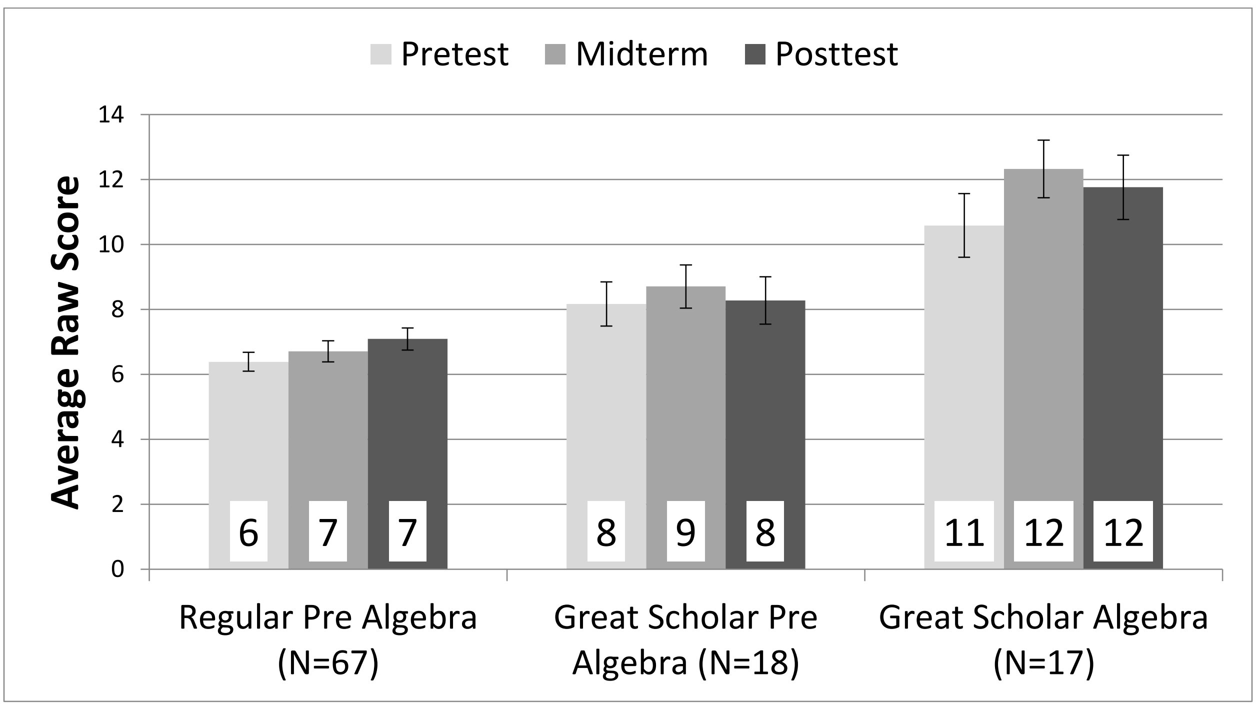


Figure 11. Math groups results over time on CTSR. Regular Pre-Algebra (left), Great Scholar Pre-Algebra (center), and Great Scholar Algebra (right) average raw scores on pretest (light gray), midterm (medium gray), and posttest (dark gray).

had an average pretest score of (8.2 ± 0.7), an average midterm score of (8.7 ± 0.7), and an average posttest score of (8.3 ± 0.7). An ANOVA on the scores for the GS pre-algebra groups

confirms that the group saw no statistical gain (p=0.84). The GS algebra group had an average pretest score of (10.6 ± 1.0), an average midterm score were (12.3 ± 0.9), and an average posttest score were (11.8 ± 1.0). Statistical analysis using an ANOVA showed that the scores for this group was not significant (p = 0.43), indicating that the math classes did not have any statistical effect on students' abilities in scientific reasoning.

Math Groups Results on MSGI by Graphical Language

The Middle School Graphing Inventory incorporates six different graphical languages: each of the languages has certain characteristics that help embed visual information. Figure 12 below gives examples of each of the six types of graphical languages. The MSGI test had

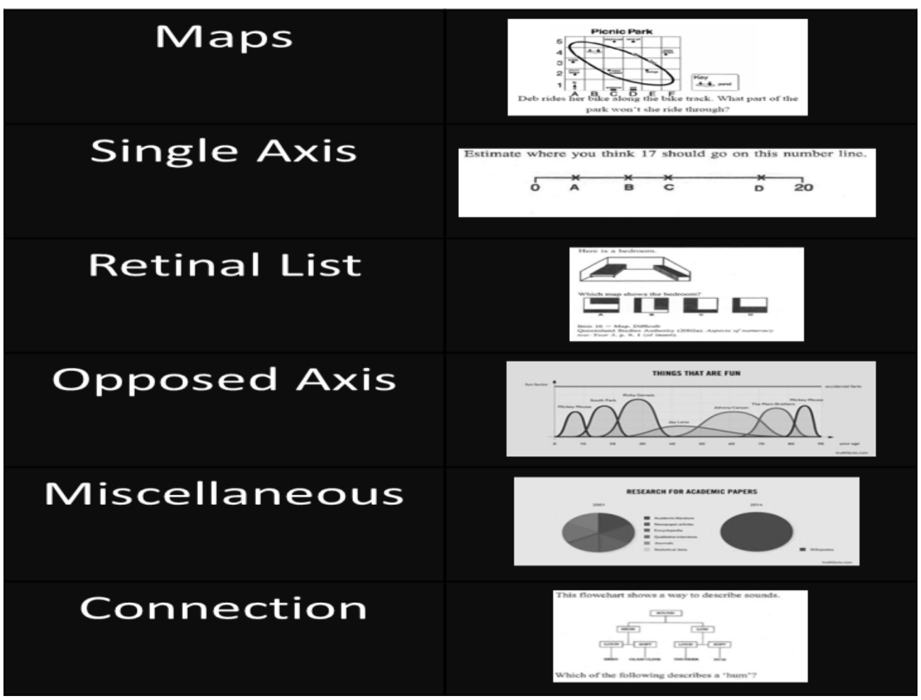


Figure 12. Examples of questions found in each graphical language

questions from each of the six graphical language categories on the test. The test had two questions in the "connection" category, four questions in the "single axis" and "miscellaneous" categories, five questions in the "map" category, and six questions in the "retinal list" and

“opposed axis” categories. The questions that were chosen for the test were primarily from the GLIM test, but were supplemented with district items. The items chosen for the test varied in complexity. Some of the items on the test were less complicated than other items and each item’s difficulty was ranked as low, medium, or high.

Since we saw that students’ pretest math ability did have an effect on their ability to read and interpret graphs, an analysis was done on each of the six types of graphical languages in each math group to see if any type of graph or groups of graphs saw significant growth. Figure 13 below shows the average pretest, midterm, and posttest scores for the regular pre-algebra group in each of the graphical languages.

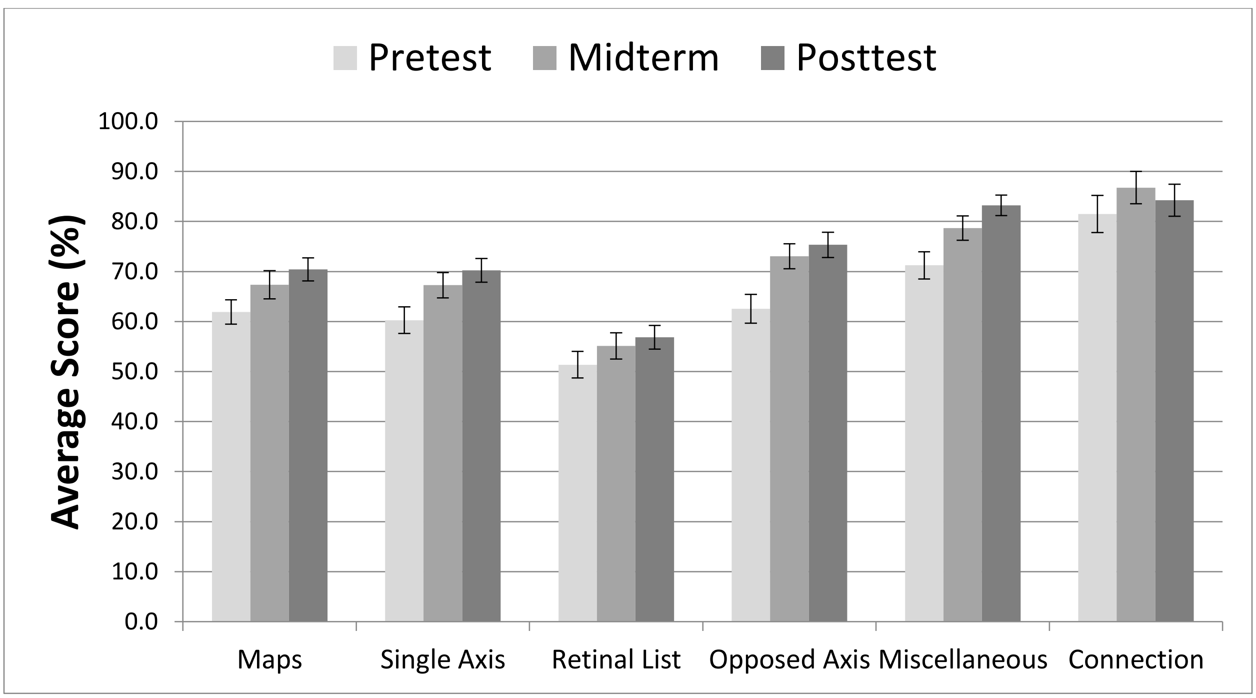


Figure 13. Regular Pre-Algebra math group (N = 73) result over time in each graphical language on MSGI. For each of the six graphical languages pretest (light gray), midterm (medium gray), and posttest (dark gray) average percentage scores are shown.

Statistical analysis was done on students’ performance in each of the graphical languages over time. ANOVAS were run on each of the groups of scores and the regular pre-

algebra groups saw significant improvements in map items (p = 0.05), single axis items (p = 0.02), opposed axis items (p = 0.001), and in miscellaneous items (p = 0.002). The items that showed significant gains were in areas that were reinforced in both the IE classes and in all of the students' math classes. The connections category and the retinal list category did not see significant growth. The reason we may not have seen a significant improvement in some of the categories is due to the small sampling of items in each of the categories and the differences in difficulty of items within each graphical language. It is worth noting that the retinal list items had the lowest average scores on the pretest and over time for all of the groups.

In Figure 14 below an analysis of the mean pretest, midterm, and posttest scores are shown and you can see from the graph that the GS pre-algebra group saw significant gains in both opposed axis (p = 0.05) and in the miscellaneous (p = 0.003). This group performs

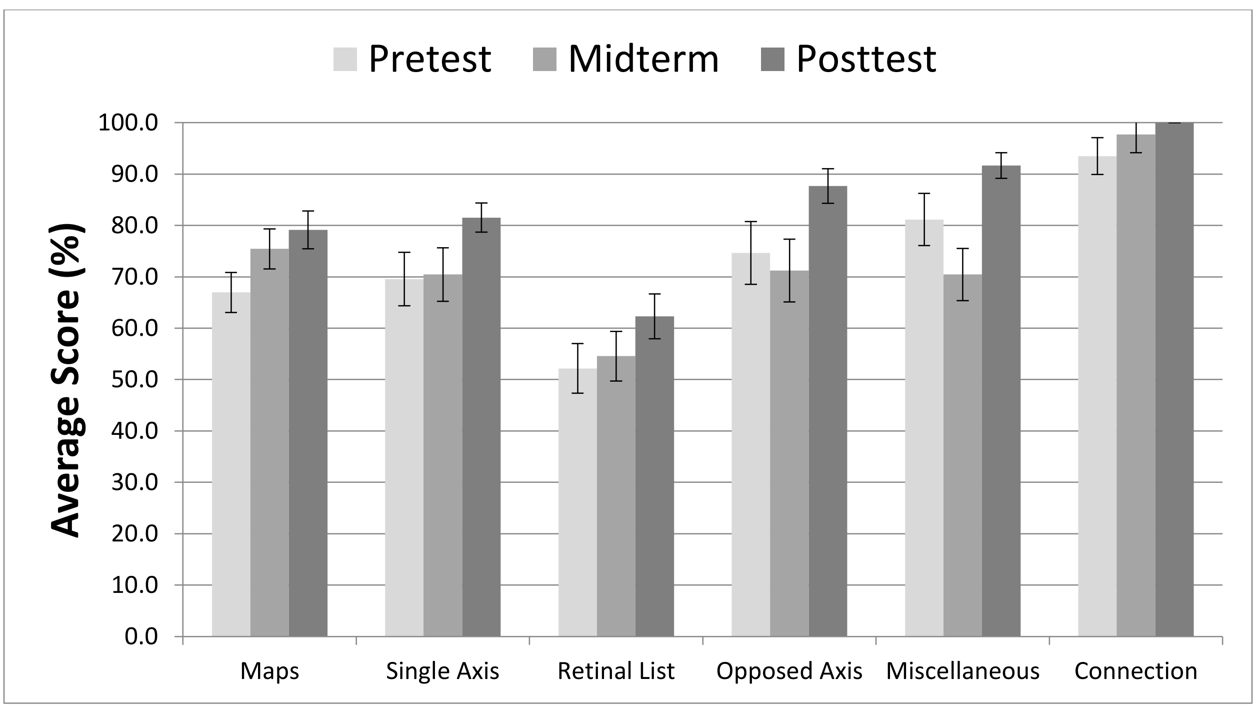


Figure 14. Great Scholar Pre-Algebra math group (N = 18) results over time in each graphical language on MSGI. For each of the six graphical languages pretest (light gray), midterm (medium gray), and posttest (dark gray) average percentage scores are shown.

similarly to the regular pre-algebra group in that we see significant growth in graphing items that were reinforced by the group's math class. Two interesting notes on the GS pre-algebra performance are that we see that they also have low average scores on retinal list items. In addition, we see the statistical growth for the GS pre-algebra group comes during the second semester when they were doing graphing items in their math classes, but not during the first semester when some of these students were in the science classroom that was utilizing IE strategies.

In Figure 15 below, we see the results for the GS algebra group over time in each of the graphical languages and an ANOVA on each of the scores showed that the algebra students did not see statistical gain in any of the groups. A possible reason for this is that they had higher average scores than the other two groups, indicating higher graphing ability. This group also had significantly higher pretest average scores and we see that there scores remained high, but increased less than the other two math groups. This information, combined with the knowledge that math classes do affect students' abilities in graphing, might indicate that the students in this group had a high level of skill because of their math abilities. This would explain why we did not see gains in the algebra group because the students were already graphically proficient.

When analyzing math groups' performance in graphing ability and science reasoning, we found several interesting results. We saw the particular math background they had did not affect their science reasoning. Pre-algebra students improved significantly in their ability to interpret graphical information due to both IE strategies in the classroom and pre-algebra's focus on graphs. In contrast, algebra students did not see any significant changes in graphing ability as measured by the MSGI.

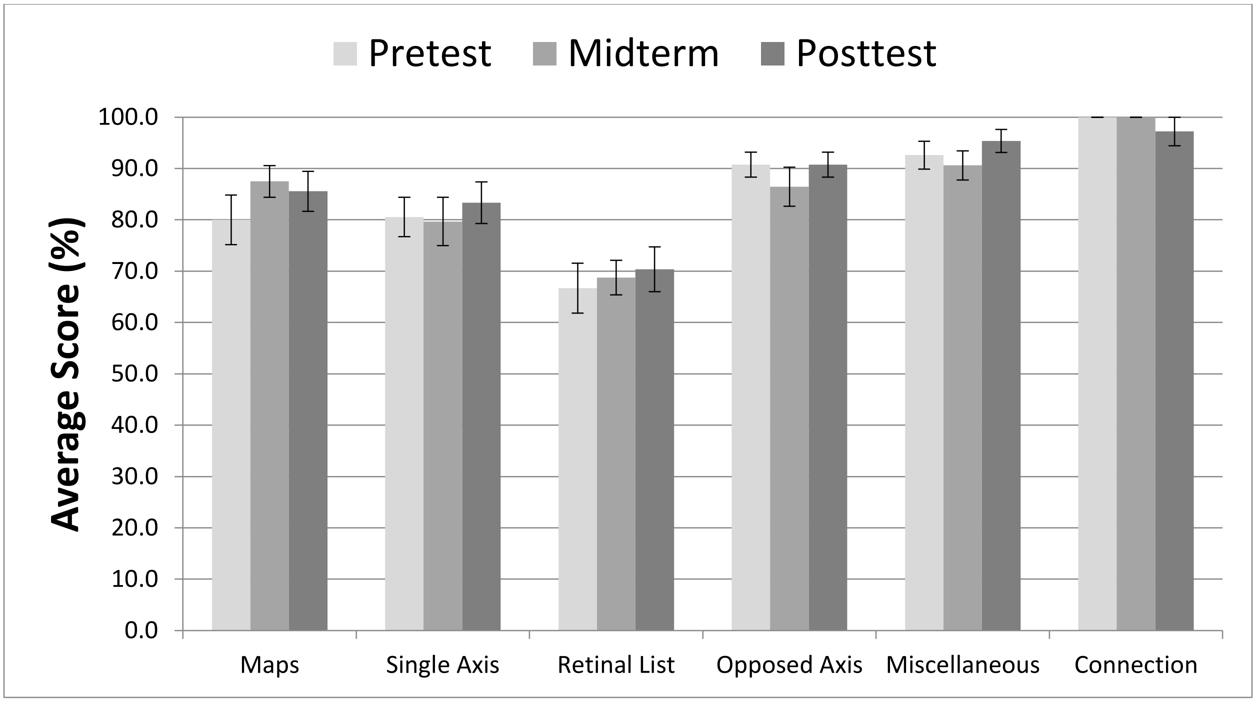


Figure 15. Great Scholar Algebra math group (N = 17) results over time in each graphical language on MSGI. For each of the six graphical languages pretest (light gray), midterm (medium gray), and posttest (dark gray) average percentage scores are shown.

When we look at performance in each graphical language, we saw improvements over time in pre-algebra students, particularly in the areas of opposed axis and miscellaneous items. The connection item category consistently saw the best performance in all math groups, while performance was consistently the worst in the retinal list items.

These results may indicate that students are best at connection items and worst in retinal list items, but it could also indicate general weaknesses in the test in that the item difficulty within each category was not equal. The retinal list items obviously caused more problems to students than the other groups. Two possible reasons for the difficulty were that retinal list items required some skill at spatial awareness or possibly just that the retinal list items were more difficult. The retinal list items chosen were mostly rated as medium difficulty and high difficulty. We also see that students on average did not find the connection items

difficult. The test only asked two questions in this category however and both of the questions were rated as low to medium difficulty. Since this test is still under verification as a valid assessment tool, it is important to note the concerns and limitations to the test before drawing conclusions from its results.

CONCLUSIONS

This study examined the impact of student centered learning strategies in a middle school earth science classroom. The existing research paints a clear picture that many science disciplines have seen significant conceptual gains in the classroom when effective instructional methods are implemented in the classroom that allow a student to learn science in the way science was meant to be learned: through exploration. The research from this study suggests that instructional style does have an effect on students' ability to read and interpret graphs, but it also shows that instructional style may affect a student's ability to use scientific reasoning skills. This is in contrast to previous research using the CTSR where instructional style did not have a significant impact on scientific reasoning. Much of the previous research with the CTSR used sample groups from high school and college that had higher average pretest scores. This may indicate that instructional style can impact scientific reasoning if the intervention is given when science reasoning is low.

There is also evidence from this study that math ability plays an important role in student's ability to understand graphs. Our results indicate that math and science can both have significant impact on student learning gains in graphical understanding. Our results also show instructional method can affect a student's science reasoning in a middle school classroom. The limitations to the study were obvious from the very beginning and the results reinforce the problem: a modeling framework or similar interactive engagement strategy needs to be developed for middle school science classrooms. While gains in graphing ability did come from using IE methods, significant gains in students' knowledge of earth science will likely require creating a complete earth science curriculum using an IE framework.

REFERENCES

Alouf, J. L., & Bentley, M. L. (2003). Assessing the Impact of Inquiry-Based Science Teaching in Professional Development Activities, PK-12.

Ainsworth, S. E., Wood, D. J., & Bibby, P. A. (1997, August). Evaluating principles for multi-representational learning environments. In *7th European Conference for Research on Learning and Instruction, Athens, Greece*.

Baker, R. S., Corbett, A. T., & Koedinger, K. R. (2001). Toward a model of learning data representations.

Capps, D. K., & Crawford, B. A. (2013). Inquiry-Based Instruction and Teaching About Nature of Science: Are They Happening?. *Journal of Science Teacher Education*, *24*(3), 497-526.

Case, R., Okamoto, Y., Griffin, S., McKeough, A., Bleiker, C., Henderson, B., ... & Keating, D. P. (1996). The role of central conceptual structures in the development of children's thought. *Monographs of the society for research in child development*, i-295.

Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *AAHE bulletin*, *3*, 7.

Cleveland, W. S., & McGill, R. (1984) Graphical perception: theory, experimentation, and application to the development of graphical methods, *Journal of the American Statistical Association*, 79(387), 531-554.

Crews, K. (2006)Effects of an Inquiry-Based Instructional Model on Sixth-Grade Science Achievement, Interest, and Engagement.

Deakin, J. (2006). Ninth Grade Physical Science Students’ Achievements in Math Using a Modeling Physical Science Curriculum.

Demetriou, A., Christou, C., Spanoudis, G., & Platsidou, M. (2002). I. Introduction. *Monographs of the Society for Research in Child Development*, *67*(1), 1-38.

Diezmann, C. M., & Lowrie, T. (2008). The role of information graphics in mathematical proficiency.

Diezmann, C. M., & Lowrie, T. (2008). Assessing primary students’ knowledge of maps.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (2007). Taking science to school. *Learning and teaching science in grades K-8. Washington, DC: National Academies Press*.

Elia, I., Gagatsis, A., & Demetriou, A. (2007). The effects of different modes of representation on the solution of one-step additive problems. *Learning and Instruction*, *17*(6), 658-672.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American journal of Physics*, *66*(1), 64-74.

Halloun, I. A., & Hestenes, D. (1985). The initial knowledge state of college physics students. *American journal of Physics*, *53*(11), 1043-1055.

Harris, R. L. (1999). *Information graphics: A comprehensive illustrated reference*. Oxford University Press.

Johnson, M. A., & Lawson, A. E. (1998). What are the relative effects of reasoning ability and prior knowledge on biology achievement in expository and inquiry classes?. *Journal of research in science teaching*, *35*(1), 89-103.

Laws, P., Sokoloff, D., & Thornton, R. (1999). Promoting active learning using the results of physics education research. *UniServe Science News*, *13*, 14-19.

Lawson, A. E. (1978). The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching*, *15*(1), 11-24.

Lawson, A. E., Alkhoury, S., Benford, R., Clark, B. R., & Falconer, K. A. (2000). What kinds of scientific concepts exist? Concept construction and intellectual development in college biology. *Journal of Research in Science Teaching*, *37*(9), 996-1018.

Lawson, A. E., Banks, D. L., & Logvin, M. (2007). Self-efficacy, reasoning ability, and achievement in college biology. *Journal of Research in Science Teaching*, *44*(5), 706-724.

Liben, L. S., & Downs, R. M. (1997). Can–ism and Can'tianism: A Straw Child. *Annals of the Association of American Geographers*, *87*(1), 159-167.

Lord, T., & Orkwiszewski, T. (2006). Moving from didactic to inquiry-based instruction in a science laboratory. *The American Biology Teacher*, *68*(6), 342-345.

Lowrie, T., & Diezmann, C. M. (2007). Solving graphics problems: Student performance in junior grades. *The Journal of Educational Research*, *100*(6), 369-378.

Mackinlay, J. (1999) Automating the design of graphical presentations of relational information, in: S. K. Card, J. D. Mackinlay, & B. Schneiderman (Eds), *Readings in information visualization: Using vision to think*, (San Francisco, CA, Morgan Kaufmann), 66-81.

Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2001). *Classroom instruction that works* (Vol. 5). Alexandria, VA: Association for Supervision and Curriculum Development.

McDermott, L. C., & Redish, E. F. (1999). Resource letter: PER-1: Physics education research. *American journal of physics*, *67*(9), 755-767.

Moore, J. C., & Rubbo, L. J. (2012). Scientific reasoning abilities of nonscience majors in physics-based courses. *Physical Review Special Topics-Physics Education Research*, *8*(1), 010106.

Olson, S., & Loucks-Horsley, S. (Eds.). (2000). *Inquiry and the National Science Education Standards: A guide for teaching and learning*. National Academies Press.

Piaget, J. (1964). Part I: Cognitive development in children: Piaget development and learning. *Journal of research in science teaching*, *2*(3), 176-186.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of engineering education*, *93*(3), 223-231.

Redish, E. F., Saul, J. M., & Steinberg, R. N. (2000). On the Effectiveness of Active-Engagement Microcomputer-Based Laboratories.

Rivet, A. E., & Kastens, K. A. (2012). Developing a construct-based assessment to examine students' analogical reasoning around physical models in Earth Science. *Journal of Research in Science Teaching*, *49*(6), 713-743.

Wainer, H. (1980). A test of graphicacy in children. *Applied Psychological Measurement*, *4*(3), 331-340.

Wells, M., Hestenes, D., & Swackhamer, G. (1995). A modeling method for high school physics instruction. *American Journal of Physics*, *63*(7), 606-619.

Willett, J. B., Singer, J. D., & Martin, N. C. (1998). The design and analysis of longitudinal studies of development and psychopathology in context: Statistical models and methodological recommendations. *Development and psychopathology*, *10*(02), 395-426.

APPENDIX A MEASURING THE MEASURING TOOL

UNIT 1.1 ACTIVITY #1 - Measuring the Measuring Tool

Materials: 1 - paper clip
1 - tongue depressor
1 - new unsharpened pencil
1 - ball point pen

PART 1:

1. Collect one of each of the measuring devices listed above.

2. Measure the length of each of the following with each of the measuring devices provided.
   a. Measure the distance between the front two legs of your desk with each of the measuring devices provided. Enter the results in the data table provided. Use labels appropriate for each measurement.
   b. Measure the longest part of your desktop with each of the measuring devices provided. Enter the results in the data table provided. Use labels appropriate for each measurement.
   c. Measure the length of a paper clip with each of the measuring devices provided. Enter the results in the data table provided. Use labels appropriate for each measurement.
   d. Measure the length of a tongue depressor with each of the measuring devices provided. Enter the results in the data table provided. Use labels appropriate for each measurement.
   e. Measure the length of an unsharpened pencil with each of the measuring devices provided. Enter the results in the data table provided. Use labels appropriate for each measurement.
   c. Measure the length of an ink pen with each of the measuring devices provided. Enter the results in the data table provided. Use labels appropriate for each measurement.

	Legs	Desk	Paper Clip	Tongue Depr.	Pencil	Pen
aper clip						
ongue epressor						
encil						
k pen						

QUESTIONS:

1. What information is important when communicating these measurements with students in another classroom.

____

2. How did you deal with measurements that did not come out to a whole number?

____

3. What basis did you use to determine this?

____

____

4. Is it possible to communicate your measurements with other students in other groups? What would you have to include for them to understand your measurement? Would it be possible to communicate your measurements with someone who does not speak English? Explain.

____

____

5. Suppose the door were exactly 15 ink pens tall. How could you tell how tall the door was in paper clips without actually measuring?

____

____

6. How tall is the door in paper clips without actually measuring? **Show any calculations involved.**

____

7. How tall is the door in tongue depressors without actually measuring? **Show any calculations involved.**

____

8. Measure your height in tongue depressors. Enter here _____t.d. Without measuring, how tall are you in ink pens? _____ i.p., and in paper clips? _____p.c. **Show any calculations involved.**

APPENDIX B PRECISION OF MEASURING

UNIT 1.1 ACTIVITY #2: Precision of Measuring

Materials:

1 - paper unmarked ruler (1 glug in length)
1 - paper ruler marked in tenths of glugs
1 - paper ruler marked in hundredths of glugs
1 - pair of scissors
1 - paper clip
1 - tongue depressor
1 - un-sharpened pencil
1 - ink pen

Procedure:

1. Collect one of each of the measuring devices listed above. Cut the rulers out from the paper template (if provided).

2. Measure the length of each of the following with each of the rulers provided. Enter the results in the data table provided. Use labels appropriate for each measurement.

a. the distance between the front two legs of your desk
b. the longest part of your desktop
c. the length of a paper clip
d. the length of a tongue depressor
e. the length of an unsharpened pencil
f. the length of an ink pen

	Legs	Desk	Paper Clip	Tongue Depressor	Pencil	Pen
ank ruler						
1 glug ruler						
01 glug ruler						

QUESTIONS:

1. What information is important when communicating these measurements with students in another classroom.

2. When using the unmarked ruler, how did you deal with measures that did not come out to a whole number?

3. What basis did you use to determine this? Are you certain that your answer is correct?

4. Is it possible to communicate your measures with other students in other groups? What would you have to include for them to understand your measurement?

5. Would it be possible to communicate your measurements with someone who does not speak English? Explain.

6. Which of the rulers used would have been better for measuring the length of the room? Why?

7. Which of the rulers used would have been better for measuring the length of a paperclip?

8. Which, if any, of the rulers used would be good for measuring the distance between cities? Explain.

9. Which of the rulers would you use to measure your height to the nearest 0.01 glug? Why did you choose this ruler?

10. Use the best ruler to measure your height to the nearest 0.01-glug. _________glug

APPENDIX C WINGSPAN

UNIT 1.2: ACTIVITY #1: WINGSPAN

Your **wingspan** is defined as the distance between your left fingertip and right fingertip if you hold your arms out straight horizontally. In this activity we will be looking at how your wingspan compares with your height.

Purpose: The purpose of this activity is to:

Gain more practice in measuring length
Practice representing data in tables
Learn how to graph data from a table
Look for patterns in your graph

Materials: Each lab team should have:

A meter stick
Graph paper
String

Procedure:

1. Measure the height of each member of your lab team using the string. Then measure the length of the string in cm with the meter stick. Record this in the data table. Label the column appropriately, including units.
2. Measure the wingspan of each member of your lab team in cm the same way. Record this in the data table. Label the column appropriately, including units.
3. Now calculate wingspan divided by height, and record this in the appropriate column. Label the column appropriately.
4. After taking data and recording it in the table, one member of the group should copy the data onto the board. We want to look at the data from the entire class.
5. Proceed to the "Evaluation of Data" section to learn how to create a graph with this data.

Data:

Group Member			

Evaluation of Data:

Now that we have represented our data in a table, we are ready to learn how to create a graph to represent this data.

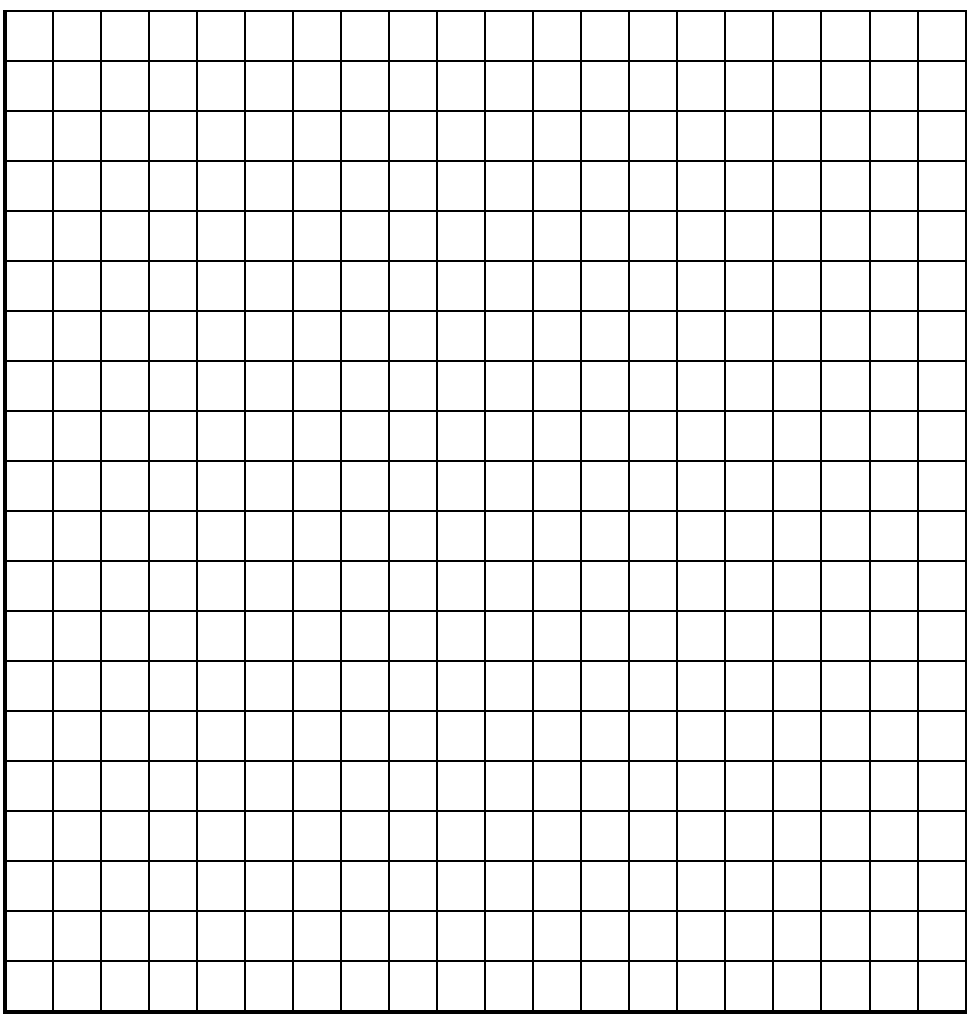
1. Label each axis of your graph. We want height (in cm) on the horizontal axis, and wingspan (in cm) on the vertical axis.

2. Mark the scale of the graph along both axes.

3. Plot the data.

4. Sketch the “best fit line.”

wingspan vs height graph



Questions:

1. Look back at your table of data. Do you see any pattern(s) emerging? If so, what pattern(s) do you see?

2. Where does your best-fit line cross one of your axes? What does this mean?

3. Where would you expect the line to cross the axes? Why?

4. Describe how you could use a graph to predict a person's wingspan.

5. An average 3 year-old is about 91 cm tall. Use your best-fit line to predict his wingspan.

APPENDIX D IRB APPROVAL

Application for Exemption from Institutional Oversight

Institutional Review Board
Dr. Robert Mathews, Chair
131 David Boyd Hall
Baton Rouge, LA 70803
P: 225.578.8692
F: 225.578.5983
irb@lsu.edu
lsu.edu/irb

Unless qualified as meeting the specific criteria for exemption from Institutional Review Board (IRB) oversight, ALL LSU research/ projects using living humans as subjects, or samples, or data obtained from humans, directly or indirectly, with or without their consent, must be approved or exempted in advance by the LSU IRB. This Form helps the PI determine if a project may be exempted, and is used to request an exemption.

– Applicant, Please fill out the application in its entirety and include the completed application as well as parts A-F, listed below, when submitting to the IRB. Once the application is completed, please the completed application to the IRB Office or to a member of the Human Subjects Screening Committee. Members of this committee can be found at http://research.lsu.edu/CompliancePoliciesProcedures/InstitutionalReviewBoard%28IRB%29/item24737.html

– A Complete Application Includes All of the Following:
- **(A)** A copy of this completed form and a copy of parts B thru F.
- **(B)** A brief project description (adequate to evaluate risks to subjects and to explain your responses to Parts 1&2)
- **(C)** Copies of all instruments to be used.
  - *If this proposal is part of a grant proposal, include a copy of the proposal and all recruitment material.
- **(D)** The consent form that you will use in the study (see part 3 for more information.)
- **(E)** Certificate of Completion of Human Subjects Protection Training for all personnel involved in the project, including students who are involved with testing or handling data, unless already on file with the IRB. Training link: (http://phrp.nihtraining.com/users/login.php)
- **(F)** IRB Security of Data Agreement: (http://research.lsu.edu/files/item26774.pdf)

1) Principal Investigator: Zane Jay Whittington **Rank:** graduate student

Dept: Natural Science **Ph:** (318)503-5473 **E-mail:** zanejaywhittington@gmail.com

2) Co Investigator(s): please include department, rank, phone and e-mail for each
*If student, please identify and name supervising professor in this space

Dr. Dana A. Browne, Professor, Dept. of Physics and Astronomy, 578-6843, phowne@lsu.edu

IRB# E8357 LSU Proposal #
- ✓ Complete Application
- ✓ Human Subjects Training
- ✓ IRB Security of Data Agreement

3) Project Title: Does modeling instruction affect learner outcomes in middle school Earth Science?

Study Exempted By:
Dr. Robert C. Mathews, Chairman
Institutional Review Board
Louisiana State University
203 B-1 David Boyd Hall
225-578-8692 | www.lsu.edu/irb
Exemption Expires: 7/14/2016

4) Proposal? (yes or no) No **If Yes, LSU Proposal Number**

Also, if YES, either ◯ **This application completely matches the scope of work in the grant**
OR ◯ **More IRB Applications will be filed later**

5) Subject pool (e.g. Psychology students) Middle School Students

*Circle any **"vulnerable populations" to be used**: (children <18; the mentally impaired, pregnant women, the ages, other). Projects with incarcerated persons cannot be exempted.

6) PI Signature Zane Jay Whittington **Date** 7/3/2013 (no per signatures)

**** I certify my responses are accurate and complete.** If the project scope or design is later changes, I will resubmit for review. I will obtain written approval from the Authorized Representative of all non-LSU institutions in which the study is conducted. I also understand that it is my responsibility to maintain copies of all consent forms at LSU for three years after completion of the study. If I leave LSU before that time the consent forms should be preserved in the Departmental Office.

Screening Committee Action: Exempted ✓ Not Exempted ____ Category/Paragraph 1

Signed Consent Waived?: Yes / No

Reviewer Mathews **Signature** Robt C Mathews **Date** 7/15/13

VITA

Zane Jay Whittington was born in Baton Rouge, Louisiana in 1979. He attended elementary, middle, and high school in Baton Rouge. Zane graduated from Scotlandville Magnet High School in May 1997. In September 1997 he entered studies at Louisiana Tech University in Ruston, Louisiana, and in March 2004 he earned a Bachelor of Science in Aviation Management. He entered the Louisiana State University Agricultural and Mechanical College in May 2011 as a candidate for the Master of Natural Sciences degree. Zane is a licensed private pilot and is currently entering his sixth year of teaching middle school earth science at Woodlawn Middle School in Baton Rouge, Louisiana.