# THE EFFECT OF SINGLE GENDER INSTRUCTION ON EIGHTH GRADE STUDENTS' MATHEMATICS ACHIEVEMENT

By

David Michael Hammel

Liberty University

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

Liberty University

July, 2013

## ACKNOWLEDGEMENTS

Throughout my educational and doctoral journey, I have so many people that I feel compelled to acknowledge and recognize for their endless support and willingness to mentor me along this path. Dr. Dozier, my superintendent, and Dr. Pringle, my assistant superintendent, have inspired, encouraged, and challenged me to grow professionally and to never be satisfied with where I am today. Without their help and support for this research project, I could never have been successful.

I also want to acknowledge and sincerely thank my doctoral committee, Doctors Yocum, Kimball, and Goggans. After only a brief meeting at an intensive summer luncheon, Dr. Yocum willingly agreed to serve as my chair and provided me with encouragement, guidance, and steady leadership for this project. Doctor Goggans' and Doctor Kimball's feedback was both insightful and effective and helped me immensely in completing this study. I am definitely a better educator and researcher because of their collective guidance, encouragement, and instruction.

I want to thank my mother and father, Patti and Mel Hammel, for always putting me first and providing me with love, support, and great role models. I am who I am today because of your steady hands and caring spirit. Words cannot express my gratitude to you both.

I also want to acknowledge and sincerely thank my wife, Julie for her endless support of my education and her willingness to always do whatever it took to ensure I could complete this journey. I would not be here without your love and encouragement and I look forward to spending the rest of my life with you and our precious children.

In closing, I want to thank God for all of the role models, friends, and family that I have been blessed with in my life and that I have so greatly influenced my personal and educational beliefs. God is truly a God of second chances and I so want to thank Him for never giving up on me.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## ABSTRACT

David Michael Hammel. THE EFFECT OF SINGLE GENDER INSTRUCTION ON EIGHTH GRADE STUDENTS' MATHEMATICS ACHIEVEMENT. Under the direction of Dr. Russell Yocum, School of Education, July, 2013.

In the research study, this investigator utilized a non-experimental, causal-comparative design (ex post facto) with archival data to determine the real impact single gender instruction had on eighth grade students' mathematics achievement. The purpose of this study was to quantitatively analyze the benefits of single gender mathematics instruction for eighth grade male and female students, when compared against traditional coeducational instruction. Specifically, the researcher compared students instructed in single gender eighth grade mathematics classes with students taught in traditional coeducational classes by analyzing fall vs. spring Measures of Academic Progress (MAP) computerized mathematics test score means. Additionally, honors vs. non-honors level groupings were analyzed. The number of students being studied (over 2,000 total from across South Carolina) and the instrumentation being utilized (Northwest Evaluation Association's (NWEA) MAP national assessment) distinguished this research project from previous studies and gave it additional validity and reliability. The mean difference in students' fall vs. spring MAP scores were analyzed utilizing a two way analysis of variance (ANOVA) to determine the impact of each main effect as well as an examination of the interaction of both effects. The research showed that the type of instruction (single gender vs. coeducational) and the interaction between the type and level of instruction did not have a significant impact on the students' overall achievement. The level of instruction (non-honors vs. honors) did have a significant

effect on the students' learning with non-honors students outperforming honors students on the MAP assessment that was administered.

# CHAPTER ONE: INTRODUCTION

## Background

The overarching goal of all educational establishments is to promote the academic success of all students. Since the beginning of formalized education, the inability to create and sustain a high level of uniform achievement for all students has plagued academic institutions throughout the United States and the world. *A Nation at Risk* (1983) brought these issues to the forefront of American consciousness almost three decades ago, and the public education system still grapples with the problem of disparate achievement levels today.

The problem of academic achievement becomes all the more serious and alarming given the overwhelming importance placed on education within today's society. With an ever changing global economy and an increased focus on highly skilled jobs and labor, a top quality, meaningful education is vital in order for the United States to retain its current position of world power, prestige, and strength. Recognizing the dire need for increased productivity and innovative techniques in America's schools, *No Child Left Behind* (2001) brought increased accountability, and the federal government relaxed many statutes and policies, such as Title IX in 2006, in order to set the stage for an innovative explosion to occur in our public education system. This type of innovation and radical change is painfully overdue and sorely needed if the United States is ever going to tackle the widening achievement gap that threatens our country's status and way of life as we know it. Currently, in order to encourage states to continue innovation and outside the box thinking even further, President Obama has relaxed *No Child Left Behind* (2001) statutes, allowing states to apply for the opportunity to further create the kind of

individualized education their students so badly need. Boyd-Zaharias and Pate-Bain (2008) echo the need for this by explaining that in order to make meaningful gains in student achievement and lower the achievement gap, significant and sweeping changes must take place in our current system. Small changes just will not make a meaningful difference.

During the same period that the United States was experiencing increased global competition and educational accountability, medical advances were also expanding at an astronomical rate. Imaging techniques and medical tests have now confirmed ideas that many long considered common sense: male and female brains are innately different (Gurian & Stevens, 2004; Sax, 2005). Furthermore, extensive educational research during the last decade shows exciting promise for boys and girls to improve their overall achievement levels in certain specific subject areas by utilizing single gender classes and innovative gender specific, research based instructional strategies (James, 2007; James, 2009; Sax, 2005).

Nationwide, single gender schools and classes are growing exponentially. South Carolina is a leader in single gender instruction and currently has 106 schools statewide implementing some form of single gender education (South Carolina State Department of Education, 2011). With single gender education becoming a popular alternative to the traditional coeducational setting, this researcher will aim to determine the true benefit of this instructional tool and ascertain if indeed this is a strategy that needs even more widespread implementation.

Current research is non-uniform and varies widely concerning the ability of single gender instruction to make a significant difference in student achievement. Numerous

studies have contradictory results and wide variation concerning the success or failure of this initiative. The low number of subjects for many of these studies combined with the lack of a specific, nationally recognized instrument makes the overall results less meaningful. This investigator's study sought to expand on the current research by utilizing 2,079 subjects in twelve South Carolina middle schools, while employing a nationally recognized standardized assessment with high levels of validity and reliability. NWEA's MAP test is nationally recognized for its outstanding validity and reliability (Northwest Evaluation Association, 2004). The use of this assessment, combined with a large number of subjects, greatly strengthened the research and provided solid evidence upon which to base any and all conclusions (Gall, Gall, & Borg, 2007).

**Problem Statement**

The problem underpinning the research study was the lack of consistent and uniform student academic achievement in America's schools today and the achievement gap that had become prevalent based on socioeconomic status, race, and gender.

**Purpose Statement**

The purpose of the research study was to quantitatively analyze the extent to which single gender instruction increases the academic achievement of eighth grade males and females in mathematics, when compared against traditional coeducational instruction. Special emphasis was given to the disaggregation of the data based on academic grouping in order to further enhance the research knowledge base concerning this educational practice.

## Significance of the Study

In the field of education, nothing is more important than the academic growth and achievement of all students. Single gender instruction holds the promise of reaching populations of students that our current instructional strategies and methods have failed to reach for decades. Furthermore, this study, if proven effective, would open the door for even greater medicinal, technological, and educational research concerning gender and maturity differences in the brain and provide new and exciting insights into ways that more effectively reach all students at all educational levels, from K-12 to higher education and adult learning.

## Research Questions

**Research Question 1:** Does the type of instruction, single gender vs. coeducational, that eighth grade mathematics students receive, make a significant difference in their overall achievement, as measured by a comparison of the fall 2011 vs. spring 2012 MAP computerized mathematics test score means?

**Research Question 2:** Does the level of instruction, honors vs. non-honors, that eighth grade mathematics students receive, make a significant difference in their overall achievement, as measured by a comparison of the fall 2011 vs. spring 2012 MAP computerized mathematics test score means?

**Research Question 3:** Does the interaction between the instructional type and the instructional level make a significant difference in eighth grade mathematics students overall achievement, as measured by a comparison of the fall 2011 vs. spring 2012 MAP computerized mathematics test score means?

## Research Hypotheses

**Null Hypothesis 1:** There is no significant difference in the overall achievement, as measured by the MAP computerized mathematics assessment, of eighth grade mathematics students based upon the type of instruction, single-gender vs. coeducational, which the students receive.

**Null Hypothesis 2:** There is no significant difference in the overall achievement, as measured by the MAP computerized mathematics assessment, of eighth grade mathematics students based upon the level of instruction, honors vs. non-honors, which the students receive.

**Null Hypothesis 3:** There is no significant difference in the overall achievement, as measured by the MAP computerized mathematics assessment, of eighth grade mathematics students based upon the interaction between the instructional type and instructional level that the students participate in.

## Definitions

**Single gender instruction** – Students are instructed in some academic area (mathematics – for this study) or areas in an all male or all female classroom environment. Both genders are still grouped together in elective classes as well as at lunch, recess, etc…

**Student Choice or "Opt-out"** – Federal law mandates that children and their parents be given the choice of a coeducational option for their child if they so desire. In most cases, this is in the same school site where single gender instruction is also occurring; however, it can involve the busing of students to different school sites if needed.

**Teacher efficacy** – The ability or capacity of a teacher to produce the desired result or effect.

**Class size** – The average number of students a teacher has in his/her class during any one period of time.

**MAP Computerized Adaptive Test** – A computer based assessment that regulates the difficulty of the questions to the student's responses. If a student is successfully answering, the questions will get progressively harder. If the student is being unsuccessful, the questions will become less difficult. The questions employed by NWEA are correlated directly to the South Carolina state standards by subject and grade level.

**Testing Window** – Each school district in South Carolina employs a time-frame, during each testing season, when students can be administered the MAP Test. Three sessions are normally utilized: a fall window at the beginning of the school year, a winter window during the middle of the school year, and a spring window just before state testing commences. These windows are normally two weeks in length and vary slightly from district to district.

## Identification of Variables

In this study, the independent variable was the type of instruction that the eighth grade students received. Middle schools in South Carolina were selected that had both single gender and coeducational eighth grade mathematics instruction occurring simultaneously in order to hold constant as many extraneous variables as possible. Additionally, only students that were enrolled in either the single gender or coeducational

mathematics class for the entire duration of instruction between the fall and spring testing cycles were considered.

The dependent variable in this research project was student achievement. For the purposes of this research investigation, student achievement was operationally defined as the difference in the mean scores of students using the 2011 fall and 2012 spring testing cycles of NWEA's MAP computerized adaptive assessment.

## Assumptions and Limitations

### Assumptions

Every research study has assumptions that, although are not tested, lie at the very heart of the investigation that is taking place. This research study had three core research assumptions. First and foremost, the primary assumption of this study centered on single gender instruction itself and its overall impact on student achievement. For purposes of this study, it is assumed that the type of classroom instruction (single gender vs. coeducational) was the primary factor which impacted student academic achievement and growth.

Secondly, this research design also had a principal assumption that undergirded its foundation. By utilizing a pre-test/post-test model, an assumption was in place that all students taking these tests gave maximum effort and performed to their maximum academic capabilities during both testing sessions.

Finally, the sampling frame also had two basic assumptions regarding the population as a whole. Principally, it is assumed that South Carolina's eighth grade middle school population was similar to the middle school population in both the United States and around the world. Furthermore, and perhaps even more importantly, it was

assumed that the eight middle school populations chosen for this study were representative of the South Carolina middle school population as a whole.

**Limitations**

As with any educational research project, there were constraints and controls which limited the researcher's ability to analyze all aspects of a problem being studied. While these limitations did hinder the overall scope and quality of the study, in no way did they diminish its importance or the significance of the results.

One limitation of this study involved the lack of measurement and examination of the teacher efficacy of both the single gender and traditional coeducational teachers involved in this project. While charting the years of teaching experience was an option, in many cases this does not clearly prove a teacher's ability or motivation.

A second limitation of this inquiry was the incapability of the researcher to ensure equal and appropriate levels of professional development in single gender strategies for all teachers that were teaching in these classrooms. In this case study, superintendents and principals of all twelve schools were not willing to provide that type of access to their classrooms.

Another limitation of this study was the variation in total class size present for each type of instruction. Due to financial hardships being placed on districts and federal law requiring opt-out clauses for all students, it was impossible to exactly match the total class size in each of these settings.

An additional, similar limitation of this research project was the inability to also equally match gender and socioeconomic groups in each type of instruction. Again, because of fiscal constraints on each school district and the option of all students to

choose the type of instruction received, there was no feasible way to match these variables exactly.

A further imperfection of this investigation was the incapability of accurately disaggregating high achievers' academic growth rates from low achievers' academic improvement rates. The MAP computerized assessment being utilized was an equal interval scaled test that accurately measured the difference scores of all students. However, students with higher pre-test scores did not improve at the same rate as students with lower pre-test results. By using NWEA's growth chart, a comparison between groups as a whole was achieved and provided a clear picture of overall student achievement.

A final, minor constraint of this study was the researcher's inability to ensure that all students were tested on the exact same school day during the fall and spring testing seasons. The reasons for this limitation centered on the lack of computer stations available to schools and districts in comparison to the number of students that were being tested and the minor differences in testing windows for each participating school district. To combat this limitation, the number of instructional days between testing was calculated and the overall differences were not significant and did not impact the study.

# CHAPTER TWO: LITERATURE REVIEW

## Introduction

Student academic achievement is the driving force and foundational goal that guides educational institutions throughout the United States and around the world. In America, accountability for student success has never been higher than it is today. *No Child Left Behind* legislation in 2002 continues a multi-decade trend of increased oversight and accountability on America's public schools from the federal, state, and local levels. Educators around the country continue to look for new and improved instructional techniques and strategies to raise the growth and achievement of many hard-to-reach students and student populations.

With advances in educational and medicinal research and technology, one of the newest strategies to reemerge in the twenty-first century is single gender instruction and its perceived ability to impact both sexes in a positive way. For most of the twentieth century, public education has been primarily coeducational in structure. While some private educational settings have remained single gender, a strong push for gender equity in the mid twentieth century resulted in federal Title IX legislation that halted or pre-empted many attempts by public schools to offer single sex classes or options for students (U.S. Department of Justice, 1972). However, in the last decade, the federal government and court systems have relaxed many educational statutes previously in place and now public school educators have the opportunity to impact learners in new and more meaningful ways, including single gender instruction (U.S. Department of Education, 2006).

Single gender instruction involves separating students based on gender in order to enhance their learning environment by removing distractions and focusing on instructional strategies that have been shown to be effective for each gender. Proponents of this type of instruction believe that boys and girls bring innate learning differences to their educational settings each and every day and by recognizing these differences, educators can more successfully educate both sexes (Gurian & Stevens, 2004; James, 2007; James, 2009; Sax, 2005). As Leonard Sax (2005), a board certified family physician and Ph. D. psychologist states in his book on gender differences:

> Today we know that innate differences between girls and boys are profound. Of course, not all girls are alike and not all boys are alike. But girls and boys do differ from one another in systematic ways that should be understood and made use of, not covered up or ignored. (p. 28)

As medical advances continue to increase and more educational research is done on this instructional tool, additional hard data will be available to be evaluated by professionals and parents alike. However, current federal law in effect since 2002 gives American public school districts the option today of providing single gender classes to students, as long as there is an option for coeducational instruction as well. This is an excellent opportunity for school leaders to provide another instructional alternative to be considered by parents and students. As Hughes (2006) asserts in her article on the advantages of single gender education:

> It has now been established that there are no legal constraints keeping public schools from creating single-sex education. Public school districts should take advantage of the opportunity to provide choice of single-sex classrooms or single

> sex schools because it is beneficial to learners, particularly minorities and those in poverty, in that their learning-styles are more easily matched, their behaviors improve, and ultimately their academic performance improves. (p. 7)

As achievement gaps between subgroups continue to widen and students of all backgrounds become more difficult to reach, additional creative options and innovative strategies such as single gender education will continue to emerge to address these issues. Continued research is the only true way to effectively analyze and evaluate the myriad of strategies that are available to educators, so that informed choices can be made that best meet the needs of all students.

The research approach for this study combined effective measurement techniques with practicality and compliance with federal law and was non-experimental in nature. As a researcher, the causal-comparative research (ex post facto) design was an excellent method for comparing the achievement of two non-random groups of individuals that have been exposed to different types of academic instruction during a determined time period.

**Conceptual and Theoretical Framework**

Three complimentary theories undergirded the research in this single gender study. While there are numerous other minor models that are sure to have had some impact in the education of our children, these three major philosophies formed the backbone of this important research.

First, one of the newest theories in educational and medicinal research today centers on innate, physiological differences in the makeup of the human brain based on gender, and the impact that these differences have had on learning, inside and outside of

the classroom (Gurian & Stevens, 2004; James, 2007; James, 2009). In the late nineteen-nineties and into the early twenty-first century, technological advances in science and medicine enabled researchers to discover new information about the human body and brain at astronomical rates. The results of this research, combined with the everyday observations of educational professionals, have had an immense impact on the way in which learning is now viewed and understood. As Gurian and Stevens (2004) asserted in their article on the differences in the male and female brain:

> New positron emission tomography (PET) and MRI technologies enable us to look inside the brains of boys and girls, where we find structural and functional differences that profoundly affect human learning. These gender differences in the brain are corroborated in males and females throughout the world and do not differ significantly across cultures. (p. 22)

However, the educational community, and society in general, have been slow to recognize this research and accept the possibility that there could be a more efficient, effective method to educate both sexes. Gurian and Stevens (2004) continued to expound on their argument by stating:

> New brain imaging technologies confirm that genetically templated brain patterning by gender plays a far larger role than we realized. Research into gender and education reveals a mismatch between many of our boys' and girls' learning brains and the institutions empowered to teach our children. (p. 22)

Finally, Sax (2005) added to this body of knowledge and argument by explaining the basis for many different learning and physiological differences that have occurred based on gender. In his book he explained:

> This new research shows that females get more from their X chromosomes than males do, and that the Y chromosome in men is directly responsible for differences in the brain. Differences. Not one better than the other. Not one worse than the other. Just different. (p. 15)

This quote underlies the challenges that single gender proponents have faced from various groups ranging from the American Civil Liberties Union to the National Association of Women. While research on brain based differences is indisputable, an age old competition between the sexes for status and authority has slowed the implementation of sound educational theory in our classrooms. This research aspired to provide further sound evidence that single gender instruction has profoundly impacted the educational achievement of all students, and by doing so, to offer them additional avenues to be successful.

Secondly, extensive research has been done on closing the socioeconomic achievement gap, and many theories exist as to the reasons for the disparity and ways in which educators can successfully bridge this divide. However, little success has been achieved. Boyd-Zaharias and Pate-Bain (2008) described the dilemma in their writing on the socioeconomic and racial achievement gap:

> Low achievement and high dropout rates among poor and minority students continue to plague U.S. society. And we say 'plague" purposefully, because these children are all our children, and our nation will profit by or pay for whatever they become. While much attention over the past quarter century has focused on reforming the schools these students attend, little or no progress has been made in actually closing the achievement gaps or reducing the number of dropouts. (p. 40)

In order to effectively close the gap that currently exists between high and low socioeconomic learners, it is crucial that radical, research-based approaches be implemented. Major disparities in achievement among minorities and socioeconomic groups are present and major radical changes must take place to close these gaps. Educators have tried for decades to solve this problem with minor, insignificant enhancements of the status quo to no avail. Boyd-Zaharias and Pate-Bain (2008) went on to contend, "As advocates for equal opportunity, we must insist on transformational change. Incremental change that merely nibbles around the edges of long-term problems will fall woefully short – again" (p. 44). Single gender instruction is one such strategy that holds the promise of eradicating the present gaps that exist. Combined with other effective research-based tools, the socioeconomic achievement gap can hopefully become a memory of our recent past. This study recognized the importance of utilizing research to investigate single gender instruction and the possible impact it holds for closing the achievement gap that exists between these groups of students.

Finally, very recent research has focused on teacher buy-in and professional development and the ability of these factors to collectively impact schools and student achievement in a positive manner (Nielsen, Barry, & Staab, 2008). In past decades, teachers were more isolated and seen by many as mere implementers of higher level decision-making; however, in today's educational culture and climate, teachers want and need to be aware and involved in the decision-making process. Classroom teachers can provide valuable insight into the possibilities and pitfalls of any reform initiative, including single-gender education, and without teachers' acceptance of the program or initiative, it is very difficult to implement successfully (Turnbull, 2002). Educational

research in the last decade has tied these two vital components together. Turnbull (2002) asserted in her study on teacher buy-in to a statewide initiative that a number of factors impacted teacher decisions, but the many facets of professional development were at the heart of teacher concerns. In her article, Turnbull stated:

> Teachers were most likely to 'buy-in' to their school reform program when they had adequate training, adequate resources, helpful support from the model developers, school-level support, administrator buy-in, and control over the reform implementation in their classrooms. (p. 248)

It is clear from this statewide study that when appropriate professional support and development was a part of the initiative, teachers were much more likely to be willing adopters. The same was true in small scale research conducted at individual sites. According to Nielsen, Barry, and Staab (2008), who studied teacher reflections on the implementation of a school wide reading program, teachers first desired job related professional development that was meaningful and program related. After thoroughly learning the new skills, these teachers were ready and able to begin the implementation process. However, training and professional development had to be on-going to meet the many diverse needs and abilities of a given group of teachers as a part of any successful implementation. It was clear from this research that teacher buy in and professional development were major components to any successful implementation of a reform initiative. This research project aimed to integrate the knowledge of these three theories into a cohesive, meaningful instructional philosophy that could be implemented successfully and benefit all students and their academic achievement, regardless of socioeconomic class.

## Review of the Literature

A comprehensive review of the literature related to single gender instruction revealed a very diverse and varied collection of studies. In order to lay the groundwork for the research project and connect the reader with the current body of knowledge, this review focused on four distinct areas with regards to single gender instruction: academic achievement, academic self concept, the social/emotional benefits, and teacher buy in and professional development.

### Single Gender Instruction and Academic Achievement

The literature was quite mixed and varied regarding single gender instruction and overall student academic achievement. The situational climate and culture, gender of the student, type of school, type of subject, and previous level of achievement were all factors which seemed to impact the achievement outcomes of single gender classes and schools. Furthermore, the majority of this research was qualitative in nature and was occurring overseas in the United Kingdom, Belgium, and the continent of Australia, which had very different school climates and cultures from our own. Recent research in the southeastern United States has shown promising results with many more positive research studies than negative results being published. Finally, almost every school involved in these research projects simply split up the sexes with little or no training, buy-in from teachers, or planning (Hoffman, Badgett, & Parker, 2008; Warrington & Younger, 2002). When training was utilized prior to implementation, more positive outcomes were realized (Parker & Rennie, 2002). With these features having supplied the foundation for the current results, there were a number of conclusions which could be drawn from the current research.

First and foremost, recent research seemed to contradict earlier studies by showing that girls performed overall academically higher in single gender classes and schools. In Wong, Lam, and Ho (2002), a longitudinal study of more than 45,000 students from Hong Kong, girls performed higher in single sex classrooms and institutions, when compared against their coeducational peers. Spielhofer, Benton, and Schagen (2004) reiterated these findings:

> In contrast with such previous studies, our analysis revealed significant differences between girls in single-sex and coeducational comprehensive schools. Thus, the performance of girls in single-sex schools was a little better for almost every attainment outcome in comparison with their peers in mixed schools. (p. 149)

At the current state of implementation around the world, most research clearly showed that girls were benefitting much more from single gender instruction than their male counterparts.

As previously mentioned, the same could not be said for boys and single gender instruction. In Mulholland, Hansen, and Kaminski (2004), girls in single gender classes outperformed boys in these same classes in all areas. Furthermore, research from Spielhofer, Benton, and Schagen (2004) in England found only boys from selective schools profited academically from single gender instruction. Finally, Wong, Lam, and Ho (2002), a broad based research study in Hong Kong, found that boys entered the secondary school setting outperforming girls in all areas. Only five years later, the situation was completely the opposite. During the five years of secondary schooling, girls reversed this trend and upon graduation were superior in all subjects except

mathematics. Additionally, boys in this study performed much better in coeducational environments than single gender settings. The contradictory effectiveness of single gender instruction between the sexes seemed to point to different issues affecting this initiative. The lack of teacher training, support, professional development, and buy in, as well as each individual school's prevailing climate and culture could be predicting the success or failure of this initiative with both genders.

Furthermore, despite all of the negative data regarding boys and overall achievement, including single gender instruction, one area of single gender research for males was particularly promising. Research studies showed that lower academically functioning boys improved their achievement dramatically in a single gender education classroom. In Hoffman, Badgett, and Parker (2008), the authors of one of the very few American research projects on single gender instruction provided research which showed that even when every other group of students benefited from the coeducational classroom setting, low achievement males were more effectively served in the single gender environment. Furthermore, two studies from the United Kingdom, Spielhofer, Benton, and Schagen (2004) and Malacova (2007) also clearly demonstrated the benefits of single gender instruction in a non-selective school environment for males with low achievement histories. Finally, Groves (2005) linked single gender instruction to greater academic achievement for lower functioning ninth grade males in Western Canada. This type of research has shown much promise for single gender instruction and further research was needed in order to clearly identify the environment and procedures that must be implemented so as to maximize the benefits of this initiative.

Another area of achievement positively impacted by single gender instruction involved the learning environment and achievement levels of females in traditionally male dominated subjects, such as math or science. Van de Gaer, Pustjens, Van Damme, and De Munter (2004) followed more than 6000 students during their secondary educational experience in Belgium and measured their overall achievement levels in English and mathematics. The authors summarized their findings by stating, "Single-sex schools promote greater enjoyment, more positive attitudes, and better achievement in subjects that are traditionally viewed as gender inappropriate, like science and mathematics for girls" (Van de Gaer, Pustjens, Van Damme, & De Munter 2004, p. 318). Gillibrand, Robinson, Brawn, and Osborn (1999) reiterated these findings concerning female students and physics in a United Kingdom study when they stated:

> The other quantitative analyses are unequivocal in showing that: (1) increased confidence was positively correlated with better GCSE grades (2) final confidence was linked to choosing to proceed to A-level physics (3) choosing to do A-level physics was strongly associated with being in the single sex class in Cohort 1. (p. 359)

In this study, the faculty of a rural school that served 900 students in southwest England noticed an alarming trend. Female students were not taking physics classes at the advanced level, and the girls that were taking these classes were not showing the expected levels of achievement. The faculty and administration initiated this two year study to determine if single gender instruction could remedy this problem and it was highly effective.

Finally, a very recent American study of seventh grade female science students in an urban middle school in Michigan also strongly pointed to the benefits of single-gender instruction in this specific arena (Brooks, 2011). This study examined female students instructed in single-gender settings for a twelve week cycle in science and compared their performance to female students taught in the coeducational environment during the same timeframe. The single gender classes significantly outperformed their coeducational peers on the TerraNova Science Assessment (Brooks, 2011).

The fourth area of concentrated research on single gender instruction and academic achievement to be noted was currently occurring in middle schools in the southeastern United States. While the results in these locations were mixed as well, larger research projects were showing promising results. Houston (2011) studied fifteen middle schools in South Carolina that had implemented single gender instruction for one year. In this study, single gender classes in grades six through eight outperformed coeducational classes in all subjects on the state standardized test known as The Palmetto Achievement Challenge Test (PACT), with more students from single gender classes scoring proficient and advanced when compared against coeducational learners (Houston, 2011). While the differences were not significant, they did show the need to further investigate these findings in large numbers over multiple years. Smaller studies, however, were not as positive. Smith (2010) examined a small number of single-gender classes in one urban Tennessee middle school and no real differences in achievement results were reported. Furthermore, Elam (2009), conducted a similar small research project in Georgia middle schools with coeducational students outperforming their single-gender counterparts.

In summarizing this section, it is important to note the incredible variety of instructional contexts where single gender instruction has been implemented and the variety of results that have been achieved. The researcher hoped to significantly add to the body of knowledge on this subject by looking at a wide variety of middle schools in South Carolina with a large number of students. Additionally, through the use of the MAP assessment, a more detailed comparison of the two types of instruction could be obtained. Much more research is necessary to say with any degree of certainty the impact that this type of instruction ultimately has on students.

### Single Gender Instruction and Academic Self-Concept

One area of the female academic experience that has been positively impacted by single gender instruction is academic self-concept. Academic self-concept can best be described as a person's beliefs or ideas concerning his or her ability to be successful in an academic subject or overall academically. One way that this trait manifests itself is in female confidence in difficult subject areas. According to both Kessels and Hannover (2008) and Gillibrand, Robinson, Brawn, and Osborn (1999), female students' confidence and academic self-concept rose dramatically when placed into single gender physics classes. Furthermore, Gillibrand, Robinson, Brawn, and Osborn (1999) asserted that these female students' self-awareness increased as well. In closing interviews with the female students, the authors clearly showed that these students were very cognizant of the difference that single gender science made in their academic growth. Gillibrand, Robinson, Brawn, and Osborn (1999) explained by stating:

> Few of the girls wished for the single sex environment to be continued through to the A-level years, believing that by this stage their own competence and equality

had become evident. Without exception, however, they stated they would have chosen the single sex class again for their GCSE years.

From this statement, one can clearly see that in addition to being more self-confident and achieving at higher levels, these female students became more self-aware of the way in which they learn difficult material the best. This definitely benefited them and their long term learning goals.

This was not the only area, however, where females enjoyed unique benefits from single gender instruction. According to Carroll (2002), single gender schools in middle grades impacted female students' aspirations and goals in positive ways as well. Carroll (2002) described these distinctive advantages by stating, "When compared to girls in mixed-sex schools, girls in single-sex schools have higher aspirations, feel more efficacious about their academic achievements, and are more likely to set goals related to education" (p. 159). Furthermore, the author went on to describe the differences between these female students and their coeducational school counterparts: "In general, girls in mixed-sex schools were concerned about their physical appearance and social image, were less inclined to be motivated towards educational goals and reported career, freedom/autonomy, and interpersonal goals more frequently" (Carroll, 2002, p. 159). This research clearly demonstrated that single gender classes and/or schools impacted girls' long-term achievement in dynamic and meaningful ways.

Finally, research also demonstrated that subject areas where gender gaps exist with regards to academic self-concept, such as mathematics and science for girls, and English for boys, can be lowered and/or eliminated by single gender instruction. Sullivan

(2009) enlightened the reader in her study on single sex schooling and academic self-concept by stating:

> Girls at single-sex schools were less likely to see themselves as 'below-average' in math and science, and less likely to see themselves as above average in English than girls at co-educational schools. Boys at single-sex schools were more likely to see themselves as above average at English. (p. 281)

This study actually showed that single-gender schooling encouraged a "gender-atypical self concept". This and the preceding studies continued to reinforce the idea that girls particularly benefited in many ways from single gender instruction. More research into this field will allow educators to find new and creative ways to cater to boys and close the now present gender gap.

**Single Gender Instruction / Social and Emotional Benefits**

Substantial research also showed that girls benefited socially and emotionally from participation in single gender classes. According to Hoffman, Badgett, and Parker (2008), single gender female classes contributed to positive outcomes for girls in a variety of ways. The authors asserted, "Girls grouped by sex engaged in more academic risk-taking and participated more than did girls in CE algebra settings. In girls' classes, "there were high levels of teamwork, camaraderie, enthusiasm, and academic risk-taking" (Hoffman, Badgett, & Parker 2008, p. 26). Furthermore, girls themselves recognized the benefits of being separated from boys in coeducational schools. In a study of pupil perceptions by Jackson (2002), girls indicated a preference for single-sex classes in an area such as mathematics and enjoyed the "safe haven" that these classes provided. In general, girls in this study found the environment more relaxed, positive, and studious

when compared against their coeducational classes. Finally, a Belgian research project examined the role that single-sex and coeducational schools played in students' gender role-identity and perceived peer group acceptance. While boys and girls were very much unaffected by the type of school attended, one important piece of evidence did emerge. For both male and female students, assertiveness was seen as essential to high levels of peer group acceptance in early adolescence, and female students attending single sex schools proved to be much more assertive than their mixed sex school peers (Brutsaert, 2006). It was clearly apparent from these and other studies that girls made even greater gains when placed in single gender classes in strategic subject areas. As educators, we must begin to act on this research, as well as develop more research based strategies to reach our male population.

**Single Gender Instruction/Teacher Buy-In and Professional Development**

Single gender instruction, like any school reform initiative, has only been shown effective when the classroom teachers that were ultimately responsible for the implementation believed in its value and worth and had extensive and ongoing professional development to learn the methods necessary to apply these new strategies in the classroom setting. Warrington and Younger (2003) echoed this sentiment by describing the necessary conditions in their study on single gender instruction. These authors identified administration and teacher commitment as being paramount to the success of the implementation. Furthermore, on-going analysis and modification of practices were an integral part of the overall program. Other studies have shown the paramount importance of the foundations being established prior to implementation. According to Gray and Wilson (2006), the lack of success and teacher dissatisfaction

with a single gender program in one Northern Ireland school was directly tied to the implementation process. The authors explained by stating, "Nevertheless, consistencies between survey responses and comments made during interviews suggest the consultation process (65%), preparatory training for staff (71%), in-service training (65%) and support for teachers taking single-sex classes (65%) were considered inadequate" (p. 296). Gray and Wilson (2006) further asserted that teachers were not even made aware that this process would be started the following school year until the summer before, adding even more anxiety to an already stressful school climate. In numerous studies, an emphasis on pedagogy, teacher training, on-going support, and recognition of current teacher knowledge and assumptions were seen as effective predictors of success in single gender instruction (Hoffman, Badgett, & Parker 2008; Martino, Mills, and Lingard, 2005; Parker & Rennie, 2002), while a lack of training and support for teachers was a clear path to failure for the initiative and unhappiness and discontentment for the faculty members involved (Gray & Wilson, 2006). As Warrington and Younger (2001) emphasized in their research into the long-term effectiveness of single gender schools:

> It can be argued, however, that the single-sex mode of organization will only be most effective when it impinges explicitly on teachers' planning, teaching and assessment, and when it becomes an integral part of school and departmental development policies and staff induction programmes. (p. 353).

Before single gender instruction can ultimately be deemed a success or a failure, educational leaders, teachers, and researchers must have faithfully implemented with fidelity the foundations needed to firmly undergird this promising program.

**Summary**

It is apparent from the existing body of research that single gender instruction could be beneficial and effective for female students and for male students in certain situations. Clearly, girls profited from single gender science and mathematics classes, with results having showed an increased confidence, academic self-concept, and overall achievement (Kessels & Hannover, 2008; Gillibrand, Robinson, Brawn, & Osborn 1999). Females also benefited from the more relaxed, open, and studious environment that single gender classes provided and demonstrated more excitement, teamwork, and a willingness to be adventurous in their education (Hoffman, Badgett, & Parker 2008). Additionally, female students themselves expressed an observable understanding that single gender classes benefited them in many ways (Jackson, 2002). Finally, male students with previous low achievement also seemed to benefit from the single sex classroom environment (Hoffman, Badgett, & Parker 2008, Malacova, 2007).

Another important aspect of this research involved the teachers themselves and the actual implementation of single gender instruction. The body of research showed that in order to successfully implement single gender classes in a coeducational environment, a significant amount of teacher buy in, pre-training, and on-going professional development must have been present (Gray & Wilson, 2006; Hoffman, Badgett, & Parker 2008; Warrington & Younger, 2003, Warrington & Younger, 2001). Furthermore, Martino, Mills, and Lingard (2005) showed that the pedagogical strategies utilized and the gender beliefs and ideas of teachers could have been even more powerful than the single gender classes themselves. The point from this research was apparent. Single gender instruction by itself did not improve academic achievement outcomes for all

children. Proper planning and implementation were necessary to make this initiative a profitable one for students and teachers alike.

The overall lack of American studies and the need to investigate properly implemented single gender instruction programs are areas in which research could open new windows of understanding on this topic. Additionally, more research needs to be focused on each and every grade level and subject area, as students' social, emotional, and academic situations are highly dependent on these contextual factors.

This research provides insight into the academic and social/emotional learning occurring in American schools today. Having investigated the academic achievement of eighth grade students in mathematics in single gender classes and having compared these results with coeducational classrooms across the state of South Carolina, a clearer picture of the overall impact of single gender instruction has definitely emerged.

# CHAPTER THREE: METHODOLOGY

## Introduction

Single gender education is one of many instructional strategies being introduced throughout the United States, and particularly the southeast, in an attempt to close the achievement gap and raise the overall performance of all students (Boyd-Zaharias & Pate-Bain, 2008). While recent global research results are varied and mixed concerning this instructional tool, many research projects show promising gains for students being instructed in single gender classes. More research in American schools is needed to determine if single gender instruction can indeed raise the achievement level of all students, and more specifically to determine if different subgroups of students benefit more than others from this environment (Hoffman, Badgett, & Parker 2008; Malacova, 2007; Mulholland, Hansen, & Kaminski, 2004; Speilhofer, Benton, & Schagen 2004; Wong, Lam, & Ho, 2002).

The study outlined below was designed to measure the effect of single gender versus coeducational instruction on eighth grade students' academic achievement. This study took place in South Carolina, a state known for its single gender initiatives in public schools. The purpose of the study was to ascertain on a larger scale the true effect of single gender instruction on academic achievement as a whole, as well as its impact on differing ability groupings of students. This study included honors and non-honors level eighth grade mathematics classes.

The methodology chapter outlines the research plan which was followed, the exact processes that were implemented, and the analysis and evaluation of the results that followed the research. It is very important that all research be clearly defined and

articulated so that appropriate replication of research can occur when needed or desired by fellow researchers. This chapter provides the critical, transparent explanation.

The overall research design for the study was a non-experimental, causal-comparative design (ex post facto) using archival data from the 2011 – 2012 school year. Additionally, the following research questions and null hypotheses guided the study:

**Research Question 1:** Does the type of instruction, single gender vs. coeducational, that eighth grade mathematics students receive, make a significant difference in their overall achievement, as measured by a comparison of the fall 2011 vs. spring 2012 MAP computerized mathematics test score means?

**Research Question 2:** Does the level of instruction, honors vs. non-honors, that eighth grade mathematics students receive, make a significant difference in their overall achievement, as measured by a comparison of the fall 2011 vs. spring 2012 MAP computerized mathematics test score means?

**Research Question 3**: Does the interaction between the instructional type and the instructional level make a significant difference in eighth grade mathematics students overall achievement, as measured by a comparison of the fall 2011 vs. spring 2012 MAP computerized mathematics test score means?

**Null Hypothesis 1**: There is no significant difference in the overall achievement, as measured by the MAP computerized mathematics assessment, of eighth grade mathematics students based upon the type of instruction, single-gender vs. coeducational, which the students receive.

**Null Hypothesis 2:** There is no significant difference in the overall achievement, as measured by the MAP computerized mathematics assessment, of eighth grade

mathematics students based upon the level of instruction, honors vs. non-honors, which the students receive.

**Null Hypothesis 3:** There is no significant difference in the overall achievement, as measured by the MAP computerized mathematics assessment, of eighth grade mathematics students based upon the interaction between the instructional type and instructional level that the students participate in.

In closing, with single gender classes growing in popularity and appeal throughout the state of South Carolina, the researcher's aim was to provide research-based support for either the continuation and growth of this instructional strategy or the subsequent downsizing of this practice to focus on other, more effective, research-based methods.

**Participants**

The research project focused on middle school students, specifically eighth graders in twelve middle schools across the state of South Carolina. The study included 2,079 students, with 50.7 percent being female and 49.3 percent being male. Furthermore, approximately seventy-two percent of this population qualified for free and/or reduced lunch and as such were considered as low socioeconomic status for this project. The remaining twenty-eight percent of the students were considered as non-low socioeconomic status. The high percentage of low socioeconomic students in the study made looking directly at socioeconomics as a factor unfeasible. Furthermore, school districts were generally unwilling to release this aspect of their data to the researcher for examination and study. Finally, the students in the study were from a mixture of rural,

suburban, and urban settings, all present in the state of South Carolina and came from the four distinctive geographic and cultural regions of the state.

The sampling frame for this study theoretically included all schools in the United States and around the world that currently offered single gender instruction in their eighth grade mathematics classes. The smaller and more practical sampling frame for this proposal consisted of forty-eight middle schools in the state of South Carolina that currently offered single gender instruction in this manner. From that list of available schools, twelve middle schools were selected, based entirely on convenience sampling, region of the state, type of community served, and the willingness of these institutions to be a part of the research project. The result of this convenience sampling was that the population ratio of low socioeconomic students to non-low socioeconomic students was not as close as one to one as was desired. This limited the researcher's ability to study this factor; however, the benefits of completing this study and the possibilities for utilization of the research outweigh this minor inconvenience (Gall, Gall, & Borg, 2007).

**Setting**

The study was done utilizing 2011-2012 archival data from four school districts and twelve total middle schools in the state of South Carolina. These middle schools varied dramatically in overall achievement level, region in the state, type of community served, consistent administrative stability, teacher retention rate, and class size, among other unique variables (S.C. Middle School Report Card, 2011). The overall achievement measures for the schools that were involved varied as follows: one school rated at-risk, the worst rating possible, one school rated below average, nine schools classified as average, one school rated as good, and no schools in the study evaluated as excellent, the

highest rating possible. Additionally, these schools represented all four geographic and cultural regions of the state of South Carolina and served rural, suburban, and urban populations. Furthermore, the overall configuration and organizational stability differed significantly among these institutions of learning. The principal's experience for the schools in the study ranged from 1 to 9 years at the school with an overall average of 4 years experience in the current position. The ability to preserve and maintain a consistent teaching staff was also variable and fluctuated between 76% and 94% with an overall average of 84.4%. Finally, the student/teacher ratio known as class size was also highly inconsistent and uneven among the schools involved in this project. Class size ratios ranged from an average of 18 at the smallest to 28 on the upper end. While these values were lower than many nationwide, they did give the researcher a broader base from which to justify the findings (S.C. Middle School Report Cards, 2011).

In closing, these schools were specifically chosen not only because of their willingness to participate, but because they represented a wide cross section of the state of South Carolina. Through this very diverse collection of educational institutions, the researcher aimed to make this study more relevant and meaningful to the overall middle school educational community in the United States today.

**Instrumentation**

A real strength of the research proposal was the measurement tool used to calculate overall student achievement. NWEA's MAP computer based adaptive assessment, first utilized in 1977, brings with it over thirty years and twenty-four million units of data from which to nationally norm each and every child. NWEA clearly articulates the strength in validity and reliability of the MAP test:

> Test and re-test studies have consistently yielded statistically valid correlations between multiple events for the same student. Most such studies rely on the methodology of having students re-test within several days. NWEA test and re-test studies have typically looked at scores from the same students after a lapse of several months. Despite this methodology (which would have the expected result of lowering the correlation figures) the reliability indices have consistently been above what is considered statistically significant. (p. 1)

Furthermore, the researchers at NWEA go on to assert the internal reliability as well:

> Internal reliability (reliability between test items) has also been impressive. This is all the more remarkable in view of the volume and breadth of the item bank, and the fact that MAP is an adaptive test. MAP users can be confident of the reliability of their tests (p.1)

In looking at these individual studies run in 2002, eighth grade mathematics had a test-retest reliability Pearson correlation coefficient equal to .93 with a sample size of 46,425 students. Additionally, the content validity was sufficiently high, with all questions being tied directly to the South Carolina eighth grade mathematics standards and the concurrent validity Pearson correlation coefficient is .85, also satisfactory for use in this study.

Finally, Hauser and Kingsbury (2009), researchers at NWEA, have also *recently* developed a method for examining "individual score validity" based on the response time a student exhibits to each question on the computer screen. This latest research serves to only enhance the validity and reliability of NWEA's MAP test and build on its impressive reputation as an outstanding assessment tool. In closing, by using this nationally norm-referenced measurement tool, with a proven history of strong validity

and reliability, this research project was able to withstand any and all reliability and validity challenges.

**Procedures**

This study was conducted through the cooperation and support of twelve South Carolina Middle Schools, representing four diverse South Carolina school districts across the four regions of the state of South Carolina. The researcher's superintendent had a good professional relationship with each district's superintendent represented in this project. With this collegial relationship already in place, a written request was made to each superintendent, along with a personal phone call from the researcher and the researcher's superintendent. In this formal request, all data needed for this proposal was requested: eighth grade mathematics class rosters showing subject level assignment as well as 2011-2012 eighth grade mathematics students' MAP scores, including the fall (pretest) and the spring RIT score (posttest). Finally, the last piece of information needed for this study was the date that each student or class was given their fall and spring assessment. This information helped the researcher determine if a significant difference was present in the amount of instruction given to any group or set of students.

The initial step in the data collection process involved acquiring authorization from the Liberty University Institutional Review Board. Upon clearance from the Liberty University Institution Review Board, this researcher and his superintendent began the process of requesting all of the needed data. This process took approximately twelve weeks and at that time the researcher was ready to analyze and evaluate the data received.

The initial step in the evaluation of the data was the entering of the data into the statistical package for the social sciences (SPSS) software program.

The next stage in the analysis and evaluation of the data was the descriptive statistics calculation and analysis. After the descriptive statistics were computed, the researcher looked at the mean pretest scores as a whole and for all subgroups being analyzed. If the pretest mean scores were equal or very close for each subgroup, then an analysis of variance for repeated measures could be utilized. However, just as the investigator expected, the pretest mean scores were not similar. This was not surprising, due to the lack of random assignment in the study. Due to a difference in the initial means that was present, the researcher attempted to use an ANCOVA to control for these differences (Gall, Gall, & Borg, 2007). However, a violation of the homogeneity of regression slopes was present and an ANCOVA was not a viable alternative. Instead, the researcher employed the most appropriate statistically feasible measure, a two-way ANOVA, in order to effectively analyze the data (Mertler & Vannatta, 2005). A two-way ANOVA was a practical option for the researcher because the data set being utilized met the statistical assumptions of normality, homogeneity of variance, and independent observations needed to conduct this statistical test.

The final step in the examination and investigation of the data involved running the suitable statistical test for significance; a two way ANOVA. After completing this process, the researcher analyzed the results obtained to determine if single gender instruction did make a significant difference in the overall achievement of eighth grade mathematics students and if any particular subgroup more greatly benefited from either type of instruction received (Gall, Gall, & Borg, 2007).

## Research Design

The design that was employed to investigate the impact of single gender instruction on the mathematics scores of eighth graders in twelve selected South Carolina middle schools was a non-experimental, causal-comparative (ex post facto) research design. In this research design, archival data from the 2011-2012 school year was obtained for students from both single gender and coeducational eighth grade mathematics classes. These groups of students were administered a computerized MAP pretest during the first two weeks of the 2011-2012 school year, known as the fall testing window or season. After a large portion of the 2011-2012 school year was completed, students were then administered a computerized MAP posttest, during April 2012, which was known as the spring testing window or season. Sample questions from the MAP assessments can be found in Appendix A. The overall mean difference between the fall and spring scores was computed for each subgroup of students along with a comparison of each instructional method as a whole (Sample South Carolina Eighth Grade Mathematics Questions, 2011).

The rationale for the selection of this design, as opposed to the more stringent quasi-experimental or experimental models, centered on federal law and the requirements of the United States Department of Education for all state education agencies. The federal Department of Education and the Office of Civil Rights explained the new less restrictive regulations by stating:

> Under these requirements, as described in the proposed regulations, the recipient must treat male and female students in an evenhanded manner in implementing its

> objective, and it must always provide a substantially equal coeducational class or extracurricular activity in the same subject or activity. (p. 62530)

While these new guidelines relax Title IX (1972) controls on public education, an important caveat was required. Students and their families must have the ability to opt-out and a suitable coeducational environment must be provided. As a result of this, single gender instruction in all public educational settings must be voluntary and no random assignment of students was possible. Additionally, archival data and an ex post facto design allowed the researcher to sample a much broader base of data than a quasi-experimental design would offer and made the overall results more able to be generalized to a larger segment of the population.

The variables of the study were simple and easy to understand. In the research project, the independent variable was the type of classroom instruction the students received, either single gender or traditional coeducational, along with the level of instruction the students participated in, either non-honors or honors. The dependent variable was the overall group student achievement, as measured by a comparison of the mean difference between the pre-test scores and the post-test scores.

The measurement of the independent variable was extremely straightforward and very clear-cut. Before the 2011-2012 school year began, parents enrolled their child or children in either single gender or coeducational mathematics classes at the honors or non-honors level. The students that remained in the single gender and coeducational class for the entirety of the school year and received a pre-test and post-test score for mathematics on MAP were considered to have participated in the study.

The measurement of the dependent variable was based on the difference between the pre-test score and the post-test score that each student received. These mean difference values were then calculated for each student and averaged together with all other students in their chosen type and level of instruction to calculate an overall mean difference for each cohort of students.

The units of analysis in the research design were the individual students being instructed in single gender and coeducational classroom mathematics classes. Students were grouped and compared by the type of instruction received, gender, and ability level (honors vs. non-honors). The goal of these comparisons was to predict which types of students benefit most from single gender or coeducational instruction.

The point of focus for this study was the comparison of academic growth or achievement of students based on the type and level of instruction the students received. The goal of the study was to determine if there was indeed a statistically significant difference between the two types of instruction.

**Data Analysis**

The analysis of data for this proposal was centered on calculating the overall mean difference in scores between the pretest and posttest for the two instructional methods being utilized. The process began with the calculation of the overall mean achievement scores, along with the standard deviations, for each type of instruction as a whole, followed by the examination of these descriptive statistics for each subgroup individually, both gender and ability grouping (honors vs. non-honors). Through the statistical tests, which are described in detail below, the researcher aimed to determine if a mean significant difference in achievement was present, based on the type and level of

instruction given, and also to compare each subgroup's performance based on this same independent variable.

In the research project, the researcher initially had two options for calculating the statistically significant difference of the mean scores. The first available option was to employ an analysis of variance for repeated measures. This type of statistical measure theoretically could be utilized because the same assessment was given for both the pretest and the posttest measure. Furthermore, it is a strong statistical test when mean scores or change is being examined. One important stipulation must be present, however, in order for an analysis of variance for repeated measures to be used: the pretest, mean scores for each subgroup being studied must be very close or equal. Due to the lack of random assignment in the study, this qualification was problematic. Upon collection and examination of the data, the examiner determined that the mean pretest scores did not meet this vital qualification.

A second more advanced technique was also available to be utilized for this research proposal. In studies which are causal-comparative in nature and involve non-random assignment, the possibility exists that large differences are present in the pretest data by group. When this occurs, an ANCOVA can be employed to correct and account for these original dissimilarities and to then make a comparison of all groups. In order to use this technique, however, certain "assumptions must be met" and checked prior to starting. Gall, Gall, and Borg (2007) explain this intricate process when they state:

> Research data need to satisfy certain statistical assumptions before analysis of covariance can be applied. These assumptions, such as homogeneity of regression, can be checked empirically, but the computations are complex.

> Inexperienced researchers need to consult an expert statistician before using analysis of covariance. (p. 321)

In this study, an ANCOVA could not be run because of a violation of the essential assumption of homogeneity of regression slopes. Due to this violation taking place, the researcher moved to the most statistically viable test available, a two-way ANOVA (Mertler & Vannatta, 2005).

The study was conducted at an alpha significance level of .05 and with 2,079 students; the research project had more than enough subjects to ensure an appropriate level of power. A large number of members for the study were needed for an appropriate power measurement, as the effect size for this project was presumably small, based on previous research (Howell, 2008).

In this research project, a determination was made, concerning which statistical technique to utilize, upon an analysis of the mean pre-test scores and discussions with my dissertation chair and committee. Finally, all decisions that were made concerning the selection of the two way ANOVA as the statistical test centered around making this research proposal as meaningful and data driven as possible.

# CHAPTER FOUR: FINDINGS

The purpose of this research investigation was to quantitatively measure the impact that the type of instruction, single gender vs. coeducational, had on eighth grade students' performance in mathematics. Additionally, the collected data were broken down by level of instruction, honors vs. non-honors, in order to examine the impact of this practice and its interaction with the type of instruction. For this project, the researcher collected two sets of test scores for students; a fall 2011 and a spring 2012 MAP score, along with each student's gender, type, and level of instruction. The data was then entered into SPSS software and analyzed to determine the impact that the type of instruction, level of instruction, and the interaction between the two had on overall student achievement.

The findings of the study are broken down into eight components: (a) the summary descriptive statistics for the project, (b) the overall paired samples t-test for the dependent variable being studied, (c) the research questions and null hypotheses restated, (d) the statistical test utilized and its underlying assumptions, (e) the results and data examination for the two way ANOVA on the type of instruction and its impact on student achievement (f) the data analysis and outcomes for the two way ANOVA on the level of instruction and its overall significance relating to student academic success, (g) the data breakdown and results for the two way ANOVA on the interaction between the two independent variables, and (h) an overall summary of the findings for the study.

## Descriptive Statistics

In the study, data was collected from twelve middle schools representing four different school districts across the state of South Carolina. A total of 2079 students' fall

2011 and spring 2012 MAP score data were received. Table 1 displays the overall descriptive test score data from the study including the minimum and maximum scores for each setting, the means, and the standard deviations.

Table 1

*Overall Descriptive MAP Test Score Data*

| | *N* | Minimum | Maximum | Mean | Std. Deviation |
|---|---|---|---|---|---|
| Fall 2011 | 2079 | 156 | 279 | 229.38 | 15.100 |
| Spring 2012 | 2079 | 162 | 311 | 233.88 | 15.806 |

Additionally, Table 2 below displays the mean, standard deviation, and mean difference between the fall 2011 and spring 2012 MAP scores by type of instruction in accordance with the first research question that stated: Does the type of instruction, single gender vs. coeducational, that eighth grade mathematics students receive make a significant difference in their overall achievement, as measured by a comparison of the 2011 fall vs. 2012 spring MAP computerized mathematics test score means? For this portion of the data, 696 (33.5%) single gender test scores were collected and 1,383 (66.5%) coeducational test scores were confirmed and utilized.

Table 2

*Descriptive MAP Test Score Data by Type of Instruction*

| Instruction type | | Fall 2011 | Spring 2012 | Mean Difference |
|---|---|---|---|---|
| Single gender | Mean | 231.33 | 236.28 | 4.95 |
| | *N* | 696 | 696 | 696 |
| | Std. Dev. | 15.25 | 16.35 | 7.93 |
| Coeducational | Mean | 228.40 | 232.67 | 4.27 |
| | *N* | 1383 | 1383 | 1383 |
| | Std. Dev. | 14.93 | 15.39 | 7.64 |

Finally, Table 3 exhibits the mean, standard deviation, and mean difference between the fall 2011 and spring 2012 MAP scores by level of instruction as it relates to second research question that asserted: Does the level of instruction, honors vs. non-honors, that eighth grade mathematics students receive make a significant difference in their overall achievement, as measured by a comparison of the 2011 fall vs. 2012 spring MAP computerized mathematics test score means? Out of 2,079 total data scores received, 1,534 (73.8%) were non-honors students while 545 (26.2%) were honors level learners.

Table 3

*Descriptive MAP Test Score Data by Level of Instruction*

| Instruction Level | | Fall 2011 | Spring 2012 | Mean Difference |
|---|---|---|---|---|
| Non-Honors | Mean | 224.25 | 229.00 | 4.75 |
| | *N* | 1534 | 1534 | 1534 |
| | Std. Dev. | 13.01 | 13.95 | 7.81 |
| Honors | Mean | 243.81 | 247.60 | 3.79 |
| | *N* | 545 | 545 | 545 |
| | Std. Dev. | 10.23 | 12.25 | 7.52 |

**Paired Samples *t*-test on Overall Student Achievement**

Initially a paired samples *t*-test was conducted between the fall 2011 mean MAP scores and the spring 2012 mean MAP scores to determine if a statistically significant difference was present for the time before and after the types and levels of instruction had been given. The results of the paired samples *t*-test were significant, $t\ (2078) = 26.480$, $p < .0005$. These findings indicate that there is a significant increase in the overall scores of the students for the instructional period between the fall 2011 MAP test (M = 229.38, SD = 15.100, N = 2079) and the spring 2012 MAP test (M = 233.88, SD = 15.806). The overall effect size calculated using eta squared = .25 and the overall effect size is large.

**Research Questions and Hypotheses**

This study was designed to examine the impact that the type of instruction, single gender vs. coeducational, and the level of instruction, honors vs. non-honors, had on eighth grade mathematics students' overall academic achievement. Furthermore, the

researcher also studied the influence that these two variables had in interaction with one another concerning the overall student achievement in eighth grade mathematics. The following three questions and hypotheses undergirded this research.

Research Question 1: Does the type of instruction, single gender vs. coeducational, that eighth grade mathematics students receive make a significant difference in their overall achievement, as measured by a comparison of the fall 2011 vs. spring 2012 MAP computerized mathematics test score means?

Null Hypothesis 1: There will be no significant difference in the overall achievement, as measured by the MAP computerized mathematics test, of eighth grade mathematics students based upon the type of instruction, single gender vs. coeducational, which the students receive.

Research Question 2: Does the level of instruction, honors vs. non-honors, that eighth grade mathematics students receive make a significant difference in their overall achievement, as measured by a comparison of the fall 2011 vs. spring 2012 MAP computerized mathematics test score means?

Null Hypothesis 2: There will be no significant difference in the overall achievement, as measured by the MAP computerized mathematics test, of eighth grade mathematics students based upon the level of instruction, honors vs. non-honors, which the students receive.

Research Question 3: Does the interaction between the instructional type and the instructional level make a significant difference in eighth grade mathematics students overall achievement, as measured by a comparison of the fall 2011 vs. spring 2012 MAP computerized mathematics test score means?

Null Hypothesis 3: There will be no significant difference in the overall achievement, as measured by the MAP computerized mathematics test, of eighth grade mathematics students based upon the interaction between the instructional type and instructional level the students participate in.

**Statistical Test Utilized**

**Two Way ANOVA.** A two way ANOVA using the SPSS software program was employed to conclude if a significant difference was present between the mean fall 2011 MAP scores and the mean spring 2012 MAP scores based upon type of instruction, level of instruction, and the interaction between the two independent variables. The type of instruction acted as the first independent variable and had two levels: eighth grade mathematics students that received single gender instruction vs. eighth grade mathematics students that received coeducational instruction. The second independent variable in the study, level of instruction, also had two components: eighth grade mathematics students that received honors level instruction vs. eighth grade mathematics students that received instruction at the non-honors level. The dependent variable in the research study was the overall mean difference scores between the fall 2011 and spring 2012 MAP testing cycles for each group of students collectively.

***Assumptions.*** Before running the two way ANOVA, initial statistical tests were performed to ensure that all of the required assumptions were met for utilizing this statistical measure. In order to effectively employ a two way ANOVA, three statistical assumptions must be met including, (a) normality, (b) homogeneity of variance, and (c) the assumption of independent observations.

In the research study, dual tests were performed to ensure that the assumption of normality was met. The probability p-plot graph below showing expected vs. observed test score distribution for the fall 2011 MAP score data (Figure 1) exhibits a straight line indicating normality.

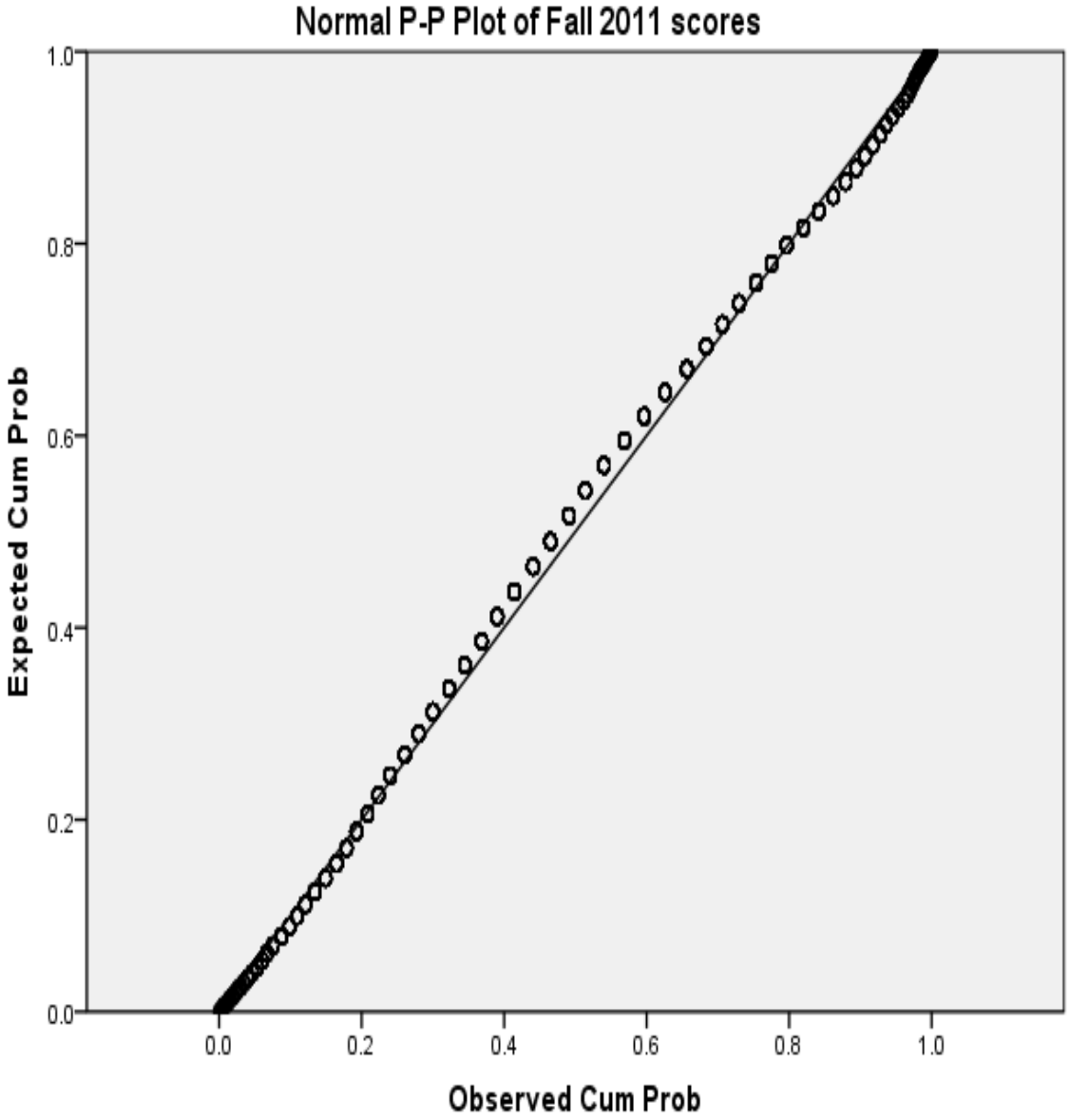


*Figure 1.* Probability-Probability Plot of Fall 2011 MAP Assessment Student Scores (Observed Vs. Expected Cumulative Probability).

Additionally, a probability p-plot graph below was also calculated for the spring 2012 MAP score data set (Figure 2) and it also clearly proves normality

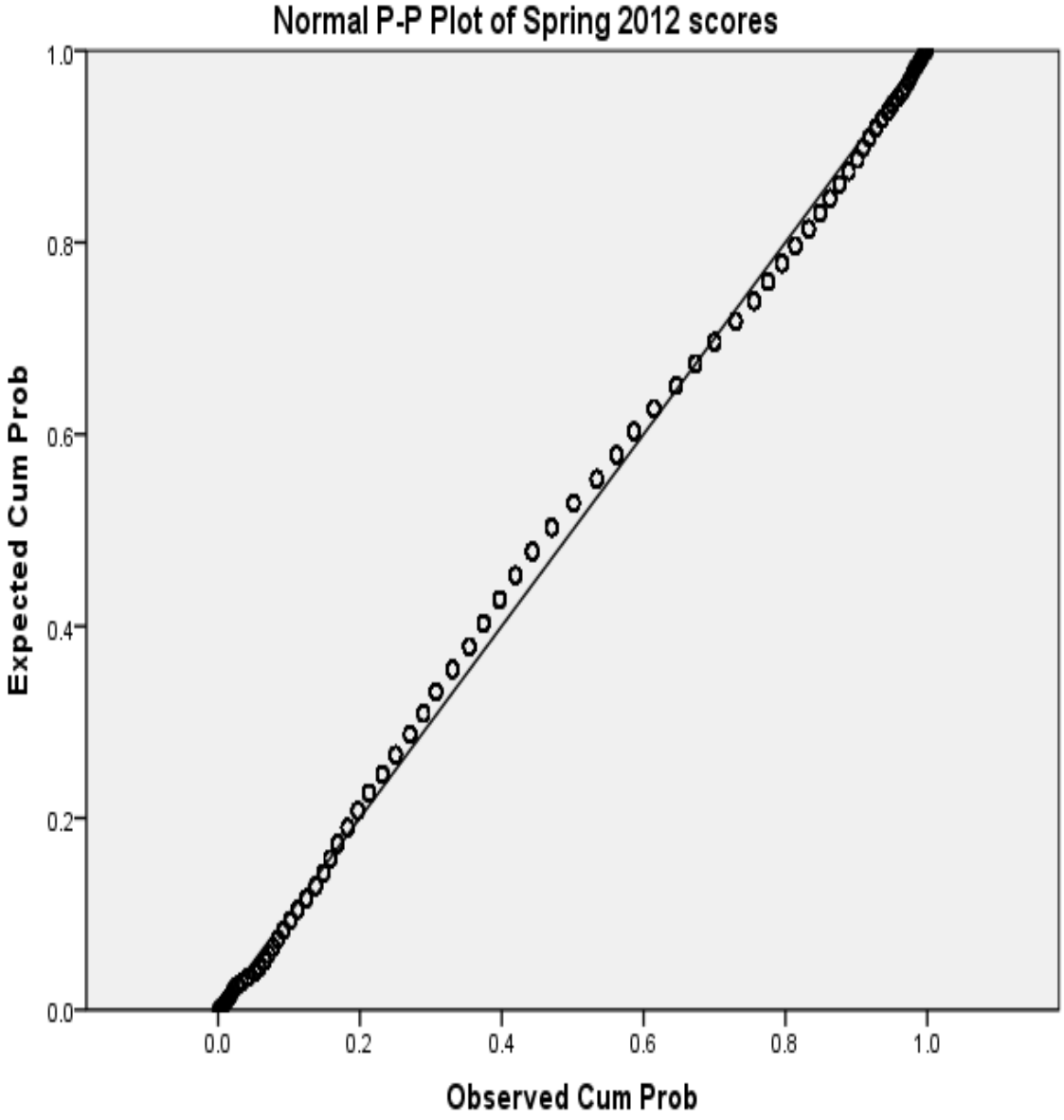


*Figure 2.* Probability-Probability Plot of Spring 2012 MAP Assessment Student Scores (Observed Vs. Expected Cumulative Probability).

Furthermore, a histogram of the fall 2011 MAP score data (Figure 3) shown below also indicates a strong normal distribution and confirms the findings of the probability p-plot graph (Figure 1).

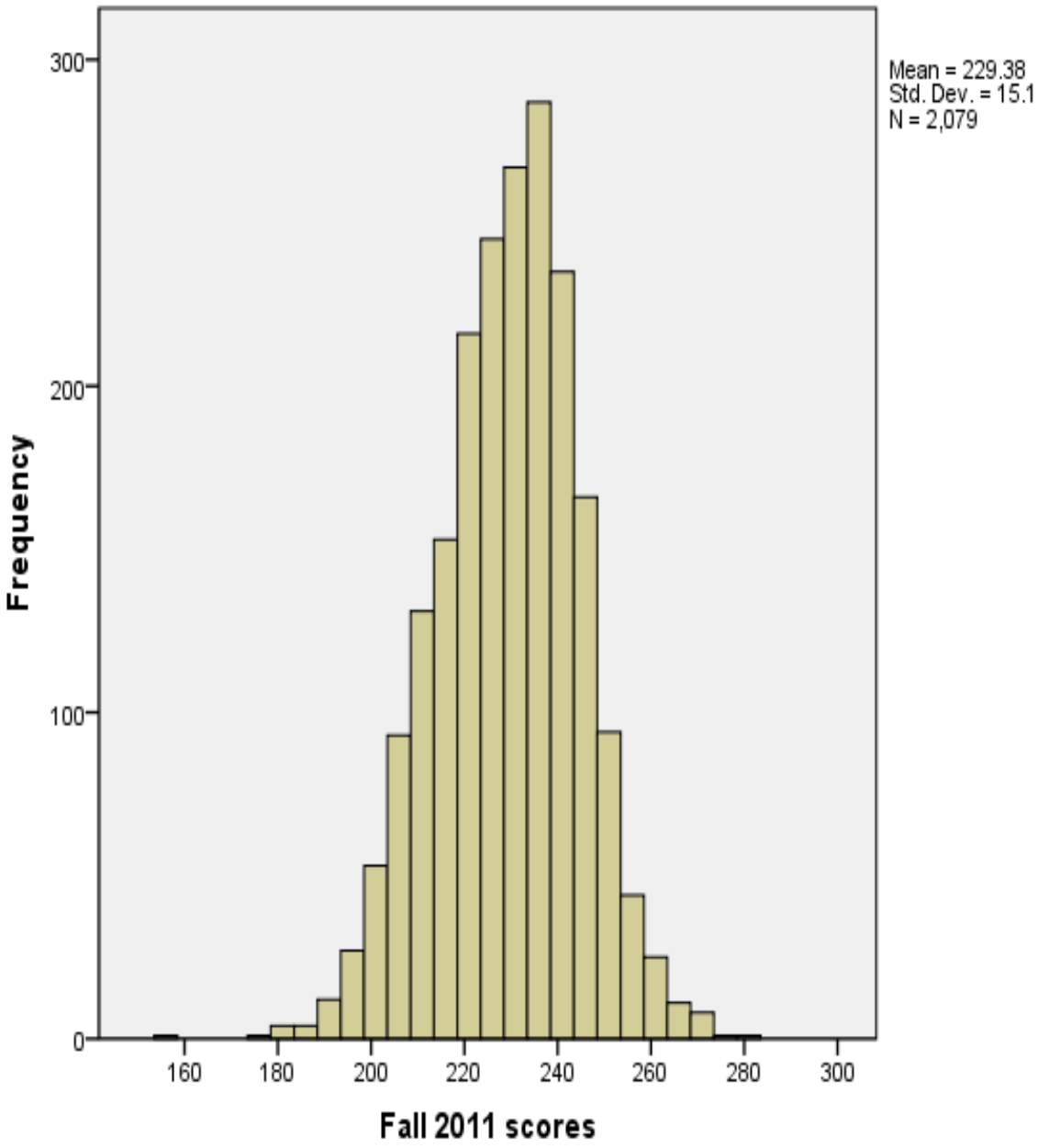


*Figure 3.* Histogram of Fall 2011 MAP Assessment Student Scores (Rausch Unit Numerical Student Score Vs. Frequency of Score)

Finally, a histogram of the spring 2012 MAP score data (Figure 4) shown below reiterates the results of the previous probability p-plot graph (Figure 2) by displaying a robust normal distribution as well.

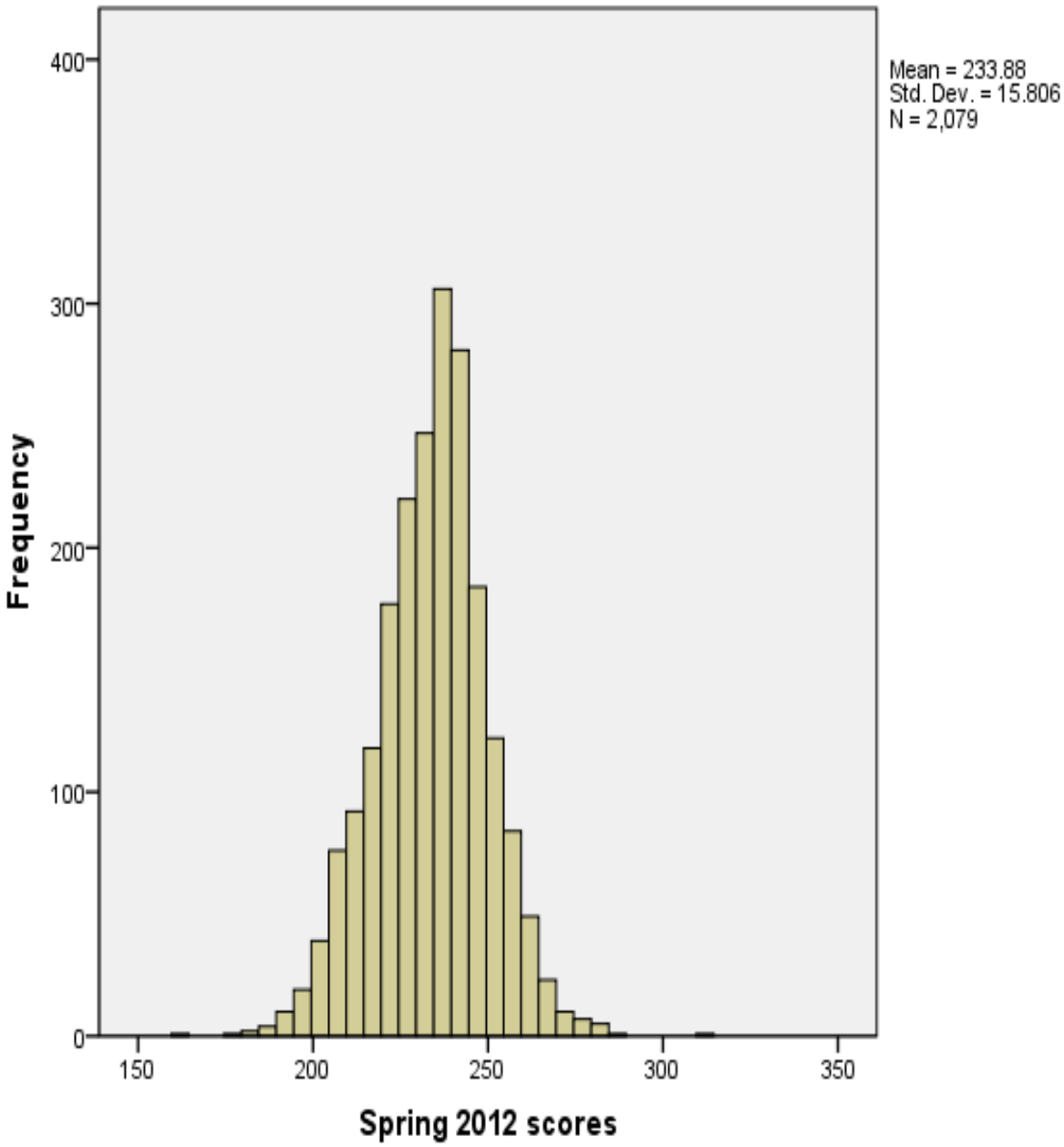


*Figure 4.* Histogram of Spring 2012 MAP Assessment Student Scores (Rausch Unit Numerical Student Score Vs. Frequency of Score

Based on the results of the probability p-plots and the histograms examined, the researcher confirmed that the assumption of normality has been met.

The second assumption that must be met in order to run a two way ANOVA is the assumption of the homogeneity of variance. Homogeneity of variance is the assumption that the dispersal of test scores for the dependent variable for all groups being studied, including the type of instruction, level of instruction, and interaction must have a consistent variance. In order to check this assumption, Levene's Test of Equality of Error Variances was conducted. The calculation of this statistical test produced an $F(3,2075) = 2.175$, $p = .091$; therefore the assumption that all of the groups being tested come from populations with equal variances is confirmed.

The third and final assumption needed in order to conduct a statistically valid ANOVA is the assumption of independent observations. This assumption is dependent upon the design of the research study and can be examined by looking at the way in which the test data was collected from the subjects. In this case, all students independently took both their fall 2011 and spring 2012 MAP computerized mathematics assessment. There was no talking or interaction between the students during the assessment period and each student's test was unique and dependent only upon the student's responses in order to generate future questions. Finally, each testing cycle, the fall 2011 and spring 2012, was independent from all previous and future administrations and in no way tied in previous student results to the current assessment being given.

**Data Analysis and Results for Type of Instruction**

An ANOVA was conducted to determine if the type of instruction, single gender vs. coeducational, that eighth grade mathematics students received made a significant difference in the students' overall achievement. The dependent variable used in measuring the achievement of the students was a comparison the fall 2011 and spring

2012 MAP computerized mathematics test score means. A total of 2,079 students' fall 2011 and spring 2012 MAP scores were examined as a part of this study. From those 2,079 students, 696 were taught using single gender instruction, comprising 33.5%, while 1383, 66.5%, were instructed in a coeducational classroom. The unbalanced design (two coeducational students for every one single gender student) was the result of the data made available to the researcher and the overall prevalence of coeducational classrooms compared with single gender classrooms throughout the state of South Carolina. The type of instruction students received, single gender vs. coeducational, was confirmed to be a non-significant factor in their overall achievement with $F(1, 2075) = 1.78$, $p < .184$, $\eta^2 = .001$. Additionally, the type of instruction only accounted for .1% of the variation seen in the test scores of the students. Finally, table 4 shows the estimated marginal means and the 95% confidence interval levels for the type of instruction students received. While single gender students' estimated marginal mean difference was .55 points higher than coeducational students, this was not a significant difference.

Table 4

*Estimated Marginal Means and 95% Confidence Intervals for Type of Instruction*

| | | | 95% *CI* | |
|---|---|---|---|---|
| Type of Instruction | *M* | *SE* | *LB* | *UB* |
| Single Gender | 4.64 | .348 | 3.96 | 5.32 |
| Coeducational | 4.09 | .232 | 3.63 | 4.54 |

*CI* = confidence interval; *LB* = lower bound; *UB* = upper bound.

From the analysis and interpretation of the ANOVA, the null hypothesis regarding the type of instruction is confirmed. There is no significant difference in the overall

achievement of eighth grade mathematics students based upon the type of instruction that they received.

**Data Analysis and Results for Level of Instruction**

An ANOVA was also employed to determine if the level of instruction, honors vs. non-honors, given to eighth grade students, made a significant difference in the students' overall mathematics achievement. The comparison of fall 2011 and spring 2012 MAP scores was analyzed and 2,079 students again made up the sample. In this analysis, 1534 students, 73.8% were educated in non-honors settings, while 545 students, 26.2%, received their instruction in honors level classrooms. The unbalanced design of 3 non-honors students for every 1 honors student was primarily due to the high level of non-honors level classes in districts where single gender instruction was being implemented as an instructional innovation. The level of instruction the students received proved to be a significant indicator of their overall achievement with $F(1, 2075) = 5.512$, $p < .020$, $\eta^2 = .003$. While the level of instruction is seen as making a significant difference in the mean difference scores of the students, only .3% of the variation in scores is explained by the level of the students' class. Furthermore, the observed power measurement of .65 leads the researcher to conclude that under the current parameters, the null hypothesis would be rejected 65% of the time that this experiment was carried out. Table 5 lays out the estimated marginal means and the 95% confidence interval levels for the level of instruction students were exposed to. In this portion of the research study, non-honors students' estimated marginal mean difference was .99 points higher than honors students, which was calculated as significant.

Table 5

*Estimated Marginal Means and 95% Confidence Intervals for Level of Instruction*

| | | | 95% *CI* | |
|---|---|---|---|---|
| Level of Instruction | *M* | *SE* | *LB* | *UB* |
| Non-Honors | 4.86 | .207 | 4.45 | 5.26 |
| Honors | 3.87 | .364 | 3.16 | 4.59 |

*CI* = confidence interval; *LB* = lower bound; *UB* = upper bound.

From the ANOVA conducted, the null hypothesis for the level of instruction and its impact on student achievement is rejected. There is a statistically significant difference in the overall achievement of eighth grade mathematics students based upon the level of instruction that the students receive.

**Data Analysis and Results for the Interaction of the Type and Level of Instruction**

An ANOVA was also performed on the data to ascertain if the interaction between the type of instruction and level of instruction made a significant difference in the overall eighth grade mathematics achievement of the students involved in the research project. All 2,079 students' fall 2011 and spring 2012 MAP scores were again used in the running of this statistical measure. The results of the ANOVA indicate that the interaction between the type of instruction and level of instruction is not significant with $F\ (1, 2075) = .138$, $p < .712$, $\eta^2 = .000$. According to the results, the interaction between these two independent variables accounted for 0% of the variation. The estimated marginal mean differences are displayed in table 6 along with the 95% confidence intervals for the interaction between the two independent variables.

Table 6

*Estimated Marginal Means and 95% Confidence Intervals for the Interaction between the Type of Instruction and the Level of Instruction*

| | | | | 95% *CI* | |
|---|---|---|---|---|---|
| Type Instruction | Level Instruction | *M* | *SE* | *LB* | *UB* |
| Single Gender | Non-Honors | 5.21 | .33 | 4.56 | 5.87 |
| | Honors | 4.08 | .61 | 2.88 | 5.27 |
| Coeducational | Non-Honors | 4.50 | .25 | 4.02 | 4.98 |
| | Honors | 3.67 | .39 | 2.90 | 4.45 |

*CI* = confidence interval; *LB* = lower bound; *UB* = upper bound.

Based on the ANOVA carried out to assess the significance of the interaction between the type of instruction and the level of instruction, no significant interaction was found. The null hypothesis concerning the interaction between the type of instruction and the level of instruction is confirmed. There is no significant interaction between these two independent variables.

**Summary of the Results**

In the research project, the investigator employed a non-experimental, causal-comparative (ex post facto) design, utilizing archival data from the 2011-2012 school year. The researcher's goal was to assess if single gender instruction made a significant difference in the overall academic achievement of eighth grade mathematics students when compared against traditional coeducational instruction. Furthermore, the grouping of students was examined in order to assess the overall practice as well as measure the interaction between the level of instruction and the type of instruction.

Based upon the data collected, the researcher was able to determine if a significant difference existed between the types of instruction, the levels of instruction, and the interaction between these two variables. After running a two way ANOVA on the type and level of instruction with the mean difference scores serving as the dependent variable, the following results were obtained.

The type of instruction, single gender vs. coeducational, was found to be a non-significant factor in the students' overall academic achievement. The null hypothesis was confirmed for the type of instruction. The level of instruction, honors vs. non-honors was discovered to make a statistically significant difference in students' overall achievement. In the study, non-honors students performed significantly better than honors students. The null hypothesis for level of instruction was rejected. Finally, the interaction between the type of instruction and level of instruction was examined. The interaction between the two variables did not make a significant difference in the overall achievement and the null hypothesis was confirmed.

Through the statistical examination of these instructional variables, the researcher was able to determine if the utilization of single gender instruction is truly justified and beneficial for its participants. Furthermore, through the breakdown of data based on academic grouping, the investigator has re-examined this instructional tool to determine if the need for this measure is still warranted. The following chapter will discuss the implications of the findings and their impact on future educational policy and practice.

# CHAPTER FIVE: DISCUSSION

In chapter four, the results and data analysis were presented concerning the significant impact of single gender instruction on overall eighth grade mathematics student achievement. A two-way ANOVA was utilized to determine if the type of instruction, single gender vs. coeducational, had a significant impact on student learning. Additionally, the same statistical test was also employed to measure the impact that grouping, honors vs. non-honors, had on eighth grade mathematics achievement as well as to quantitatively measure the significance of the interaction between the two independent variables in this project.

The overarching goal of this chapter is to examine the results of this study and to deliberate and analyze their meaning. This chapter is broken down into six sections: an overall summary of the study, a discussion of the findings and the implications in light of the relevant literature and theory, a breakdown of the study limitations, practical and methodological implications, recommendations for future research, and a conclusion.

## Summary of the Study

Consistent and harmonized academic achievement for all students continues to be the major unachieved goal for most of America's schools. The inability of the educational institution to promote constant and uniform academic growth for all students and to close the achievement gap that is present when many students enter school plagues our establishment to this very day and threatens our democratic existence as it is currently known.

The purpose of this study was to determine on a broader scale if single gender instruction held the promise of closing the achievement gap that currently exists while at

the same time more effectively educating both male and female students. Additionally, the level of instruction, honors vs. non-honors, was examined, as well as the interaction between these two independent variables, in order to ascertain each component's impact on overall student learning and achievement.

**Research question 1: type of instruction.** The major goal of the research project was to determine the impact that single gender instruction had on overall eighth grade students' mathematics achievement. Special emphasis was given to making this study a large scale venture utilizing as many districts, schools, classrooms, and individual students across the state of South Carolina as possible. By employing a large sample combined with a nationally recognized instrument, the researcher aimed to provide educators a broad look at this type of instruction and its overall impact on learning as well as to ascertain its widespread general value as a preferred method of education.

The sample for this research question consisted of 696 students instructed in single gender classrooms, accounting for 33.5% of the total students. A total of 1,383 students, the remaining 66.5%, or approximately two times as many children, were taught in coeducational settings. The resulting non-balanced design was not ideal for this research; however, due to state and federal regulations and the lack of mandatory assignment to single gender classes, the natural disparity between the types of instruction was present. All 2,079 students were administered NWEA's MAP test in the fall of 2011 and again in the spring of 2012, to measure their overall achievement during the 2011-2012 school year.

The results of the two-way ANOVA indicated that a significant difference was not present in the students' eighth grade mathematics achievement, based upon the type

of instruction, single gender vs. coeducational, that the students received. Single gender students' estimated marginal mean difference was 4.64 points between the fall 2011 and spring 2012 testing cycles, while coeducational students averaged 4.09 points during this same period. Even though single gender students performed better overall, it was not a significant distinction. Furthermore, based upon the statistical calculations performed, the type of instruction only accounted for .1% of the variation in student scores. The null hypothesis was confirmed for Research Question 1, in that there was not a statistically significant difference in eighth grade students' mathematics achievement based upon the type of instruction the students received.

**Research question 2: level of instruction.** On a secondary scale, the researcher also aimed to examine the impact of grouping or level of instruction on overall student achievement in eighth grade mathematics. While this was not the primary goal for the research project, it was the researcher's intent to study this longstanding practice to determine its influence and appropriateness for continued usage in public education. The sample for this portion of the research study consisted of 1,534 students being instructed on the non-honors level while 545 learners were educated in an honors level setting. The unequal number of students in this sample is a direct result of the relationship between the levels of instruction being administered in school districts which have chosen to implement single gender learning. In South Carolina public schools, more innovative programs and instructional techniques were being implemented in schools and districts which serve a higher percentage of lower socio-economic students that tend to be engaged in less rigorous mathematics classes as a whole. These programs have been put

into practice in an attempt to lower the achievement gap present and move students to higher levels of overall performance.

The results of the two-way ANOVA that was conducted to examine the impact that level of instruction had on students' eighth grade mathematics achievement showed a significant difference in the students' overall performance based upon the level of instruction that they received with $F(1, 2075) = 5.512$, $p<.020$, $\eta^2=.003$ Students in non-honors level classes showed an a marginal mean difference of 4.86 points between the 2011 fall and 2012 spring testing seasons, while students in honors level classes exhibited a marginal mean difference of only 3.87 points. The observed power of .65 allows the researcher to conclude that the null hypothesis would be rejected 65% of the time that this experiment is conducted. Finally, .3% of the variation in scores can be explained by the level of instruction that the students received. Based upon the data collected and the analyses performed, the null hypothesis for Research Question 2 is rejected. There is a significant difference in the students' eighth grade mathematics achievement based upon the level of instruction that they received.

**Research question 3: interaction between the type and level of instruction.** The final research question addressed the interaction between the type and level of instruction and sought to determine the impact that each had on the other. The sample for this segment of the data analysis was again unbalanced due to the overall skewed nature of the data available to be collected. For this part of the research study, 160 students were in honors level single gender classes, while 536 learners took part in non-honors single gender classes. For the coeducational portion, 385 students participated in honors level classrooms while 998 pupils were educated in a non-honors coeducational setting.

The results of the two-way ANOVA employed to determine the significance of the interaction between the type and level of instruction revealed a non-significant interaction between the two independent variables. Students in single gender, non-honors classes showed the largest gains with a marginal mean difference of 5.21 points, while their classmates in coeducational, non-honors classes presented a marginal mean difference of only 4.50 points. Additionally, single gender, honors classes exhibited a marginal mean difference of 4.08 points, while their coeducational peers showed a marginal mean difference of 3.67 points. The null hypothesis for Research Question 3 was confirmed. There is not a significant difference in eighth grade students' mathematics achievement based upon the interaction between the type of instruction and the level of instruction that the students received.

**Discussion of the Findings**

**Research question 1: Type of instruction.** The results for the primary purpose of this research project continue the confounding debate that is present concerning single gender instruction and its overall impact on student achievement. The findings in this study show that single gender instruction did not make a significant difference in the overall academic achievement of eighth grade mathematics students, when compared against traditional coeducational instruction. This study falls in line with some of the recent research done in the United States that also indicated single gender instruction did not change the overall achievement of middle school mathematics students (Smith, 2010; Spikes, 2008; Whalen, 2012). However, other studies seem to contradict this conclusion and show that in some settings and grade levels, single gender instruction did make a significant difference in the overall achievement of middle school mathematics students

(Gary, 2011; Van Zyl, 2011; Vrooman, 2009). The conflicting results obtained from the latest American studies, with even some individual projects confirming dual opposing results at different middle level grades, possibly indicate that other confounding instructional factors are present in these learning environments.

The results of this study do confirm, however, that single gender instruction does no harm academically to the participants involved and in fact could be beneficial. Students in single gender classes produced an estimated marginal mean difference of 4.64, while coeducational classes produced an estimated marginal mean difference of only 4.09. While this achievement difference was not significant statistically, it is noteworthy to report that based on NWEA's 2011 National Norming Study and each type of instruction's starting score, single gender students were expected to show a marginal mean difference of 4.29 points, while coeducational students were anticipated to produce a marginal mean difference of 4.30 points, in order to show average growth (NWEA, 2012). From these values, single gender students met the needed difference while coeducational students did not. Additionally, recent United States and international research has confirmed older international studies that show single gender classes do improve students' self-concept and open career options to females at a higher frequency than their coeducational counterparts (Gary, 2011; Schneeweis & Zweimuller, 2012; Whalen, 2012), an area that is not addressed by this research project, but beneficial to single gender students. More in-depth research is needed to understand the complexity of this instructional strategy and its true impact on student learning, motivation, and achievement.

This segment of the quantitative research project contributes to the growing body of investigative studies concerning single gender instruction by providing researchers with the first large scale study involving over 2,000 students from multiple schools and school districts across a U.S. state. Most recent research has focused on small scale studies of individual schools and districts which have implemented this instructional tool and not on a large, state-wide scale. Furthermore, this study was only the second to utilize NWEA's MAP test, a nationally recognized instrument known for its outstanding reliability and validity.

**Research question 2: Level of instruction.** The findings concerning the secondary purpose of this research project also promote a continued discussion, analysis, and debate regarding the level of instruction and the manner in which children are grouped or placed into mathematics classes. The results of this research clearly show that the level of instruction which eighth grade mathematics students receive does make a significant difference in their overall achievement. In this study non-honors students' outperformed honors students' significantly, with an estimated marginal mean difference of 4.86 points. Furthermore, this value exceeded NWEA's projected norm of 4.33 by .53 points. Additionally, honors students did not make their projected difference norm of 4.25 points, producing an estimated marginal mean difference of only 3.87 points of achievement (NWEA, 2012). This project confirms selected recent research regarding the homogeneous grouping of students and their overall superior achievement in middle level mathematics, when compared against heterogeneously grouped students, as well as the preference of students and teachers for this type of differentiated learning (Saunders, 2005; Tieso, 2000). At the same time, contradictory research exists regarding grouping

and its overall impact on student learning, achievement, and opportunities. Burris, Heubert, and Levin (2006) asserted in their research concerning middle level mathematics and heterogeneous grouping that this practice improved learning opportunities for all children while Spielhagen (2010) echoed these same sentiments concerning homogenous grouping and its detriment to lower level students and their long term achievement and goals. The results of this study seem to confront and challenge these concerns. Overall, the conflicting research and inconsistent findings again point toward other unknown factors present in the individual studies presented.

Based on the results of the current research project, three plausible factors could be affecting the study and the results that were obtained. First, in schools with higher percentages of lower-socioeconomic students, the allocation of resources is not always evenly balanced. In many schools, extra funding and teachers are placed with lower achieving non-honors students in an effort to improve their overall performance and lower the overall class size. Additionally, honors students in South Carolina eighth grade mathematics classes are not being taught eighth grade mathematics standards during the current school year. These students have already been instructed on these standards in previous grades and are being taught high school level math courses during their eighth grade term. With NWEA's MAP test specifically designed for S.C. eighth grade mathematics standards, this lack of current instruction could be significantly impacting the results. Finally, one last possible issue impacting the research results centers on the overall teacher efficacy in honors vs. non-honors classes. All schools in South Carolina are expected to meet a minimum baseline score in order to be rated at various classifications. As a result of this external accountability, many schools choose to place

their highest efficacy teachers with their lowest achieving students in an effort to meet these minimum requirements.

This portion of the study adds to the current body of research by again supplying researchers with a large scale study of over 2,000 students from multiple schools and districts across the state of South Carolina while utilizing NWEA's nationally recognized MAP assessment.

**Research question 3: Interaction between the type and level of instruction.** The results of the final portion of this study reveal the interaction between the type and level of instruction to be a non-significant factor in eighth grade students' mathematics achievement. In this study, the interaction effect was the least significant of all statistical tests employed and accounted for none of the variation in academic student achievement. An exhaustive review of the literature also disclosed no research studies on the interaction between these two independent variables, probably due to the same factors which limited the researcher in this study of these two facets of instruction. The interdependence of these two factors and the inability to completely separate them statistically forced the researcher to use a two-way ANOVA, as opposed to the more statistically preferred ANCOVA in this project. The only current research that can be reasonably tied to this study involved current single gender female middle school students with previous gifted math instruction. Spikes (2008) sought to determine the current impact of single gender vs. coeducational instruction, while taking into account the number of previous years in gifted mathematics classes. Neither the type of instruction or number of years in gifted education was found to be significant, in accordance with the results of the current study. This study adds to the current body of

research by providing readers a clear picture regarding the interaction between these two variables in the instructional model and the overall lack of significance related to this interaction.

## Implications

**Practical advice for practitioners.** Educational leaders around the country are constantly searching for innovative strategies, initiatives, and programs that can differentiate learning and provide diverse opportunities for students to be successful. The findings of this study reveal that single gender eighth grade mathematics instruction does not harm students academically and could prove to be a viable instructional alternative for desiring students if certain other educational parameters are present. In this study, single gender students scored higher than their coeducational classmates, although a significant difference in achievement was not observed. Combining the findings from this research project with other current educational studies, it is clear that single gender instruction needs certain caveats in place, as does any instructional initiative, in order to ultimately impact students academically. First and foremost, numerous research findings show that teacher buy-in, or teacher efficacy, must be present in order to impact students academically (Gray & Wilson, 2006; Mojavezi & Tamiz, 2012; Tschannen-Moran & Barr, 2004). Without this initial precursor in place, instructional initiatives have little chance of achieving success. Additionally, current research asserts that on-going professional development must also be present to ensure the any instructional initiative reaches its full academic potential and is successful on a long-term basis (Bruce, Esmonde, Ross, Dookie, & Beatty, 2010; Hoffman, Badgett, & Parker, 2008; Martino, Mills, & Lingard, 2005; Parker & Rennie, 2002; Warrington & Younger, 2001).

Teachers must be growing professionally and receive appropriate training related to the instructional plan in order to effectively reach students. Finally, effective instructional leadership from administrators in charge of the academic program must be present for real academic achievement to occur. Numerous studies show that high expectations from educational leaders and transformational leadership practices such as goal setting, development of a vision, and individual teacher support are strong predictors of teacher efficacy, commitment, and ultimately student achievement in the classroom (Nelson, 2012; Solomon, 2007, Warrington & Younger, 2003). No educational program exists in isolation, and these institutional factors are strong predictors of the success or failure of any initiative or project before it even begins, including single gender instruction. In closing, one final practical implication of this instructional initiative involves the current situation regarding school funding in America's public schools. The ability to implement a single gender instruction initiative as an option for students, without needing additional funding or teachers, makes this instructional initiative both appealing and possible for educational leaders. With more schools turning to programs that are fiscally plausible, more opportunities will be present for researchers to study and examine this instructional tool.

A second practical implication of this research study centers on the way in which students are grouped together for instruction in eighth grade middle level mathematics. Based on the broad results of over 2,000 students in twelve South Carolina middle schools and current educational research, it is clear that with appropriate alignment and differentiation of curriculum, non-honors level students can significantly benefit from being homogeneously grouped together, in order to better meet their instructional needs

(Saunders, 2005; Tieso, 2000). In this study, non-honors students significantly outperformed their honors peers, showing an estimated marginal mean difference of 4.86 points, well above the expected gain of 4.33 points by NWEA's MAP assessment, while honors students produced an estimated marginal mean difference of 3.87 points, below the 4.25 points that this group was expected to achieve (NWEA, 2012). However, research is conflicting regarding this decades' old educational practice, and educators should be cautious of using homogenous grouping on a large scale, long-term basis. Other research studies have shown that tracking or grouping is harmful to lower level achievers and can limit their opportunities to be successful in higher level classes at the secondary level (Burris, Heubert, & Levin, 2006; Spielhagen, 2010). Based on the results of this study and existing research, homogeneous grouping in eighth grade mathematics can be successful when non-honors level students are exposed to teachers with collective efficacy and high expectations and grouping is the least restrictive and of the broadest range possible (Sauders, 2005; Tieso, 2000). Finally, the lack of academic performance and overall instructional achievement by the honors level students was somewhat surprising to the researcher, however, given that the twelve middle schools in the study served a high percentage of free and reduced lunch students and were generally lower achieving schools, a lack of instructional focus on the high achievers in these buildings was very likely. While the honors students did show gains academically, a renewed focus on high expectations and rigor could benefit these students as well.

A third and final practical implication of this study, methodological in nature, concerns the decision of the researcher to examine both the type and level of instruction together in one research project. Based upon a review of the research literature, this is

the first study that investigated the interaction between single gender and coeducational instruction, while also including the level of instruction in the equation, and the impact of this was very problematic for the researcher. Due to the inter-relatedness of the two variables involved and the corresponding unbalanced data set available to be collected, the researcher was forced to modify the research questions and change the overall statistical analysis of the data in order to proceed. During the statistical analysis, a violation of the homogeneity of regression lines was observed, preventing an ANCOVA from being utilized to analyze and breakdown the data. A two-way ANOVA was selected as the best viable alternative for the researcher to study and ascertain the significance of the two independent variables as well as to scrutinize the significance of the interaction between them. In looking at the study and the statistical issues which occurred, the researcher could have employed an ANCOVA had the data set been more balanced across both variables. However, due to the low number of single gender classes in South Carolina and across the United States, as well as the federal opt out provision which must be provided for all students, the researcher would have had to trade the large number of subjects in the study in order to balance both variables involved. The researcher's aim was to examine single gender instruction on a large scale and this was still able to be accomplished using alternative statistical measures.

**Recommendations for future research.** This research project is distinctive and adds to the current body of research by providing researchers with a broad study of over 2,000 students from four school districts and twelve middle schools in the state of South Carolina, a state known for its single gender initiatives. Furthermore, its examination of the interaction and relationship between honors and non-honors level classes and the type

of instruction students receive is the first study of its kind to be completed. Additionally, this study provides researchers with further support for using homogenous grouping in select areas and situations, as well as opens the door to further research on single gender instruction and the impact that this type of instruction might have on student learning at the state-wide level. Finally, through the use of NWEA's MAP assessment, researchers are shown a clearer picture of the impact that the level and type of instruction ultimately have on eighth grade mathematics student achievement.

In regards to future research, there are a number of recommendations to be made based upon the results of this research project:

In order to clearly ascertain the real impact of single gender instruction on middle school mathematics, a longitudinal, large scale, three year examination of single gender vs. coeducational classes should be studied. While single gender students outperformed coeducational students in the current study, the difference was not significant. A three year period would provide meaningful data to show if these differences would continue to grow to the point of significance and would give researchers a much clearer picture concerning the long term effect of single gender instruction.

Another possible area of future research related to this project centers around academic grouping in middle school math classes and its true impact on academic achievement. In this study, non-honors students achieved significantly better than their honors peers. A three-year, broad based study around the state of South Carolina or other states could show if homogenous grouping is indeed a true benefit to non-honors' students or if corrections to the test data occur over a period of time. Furthermore, it would be beneficial to examine these same students' state accountability test data at the

end of the 2011-2012 school year to determine if the MAP test results observed translated into a significant difference in achievement when compared against these students' 2010-2011 state accountability test scores.

A third area for future research focuses on employing a smaller scale study of single gender vs. coeducational instruction, while controlling for competing factors such as teacher efficacy, school climate, school culture, and overall leadership in the schools involved. This could be a one or two school project designed to help more clearly delineate the impact single gender instruction has on academic achievement. By completing these smaller studies in coordination with future large-scale examinations, a true picture of each variable's importance could be realized.

A fourth theme for future research involves an investigation concerning single gender classes in coeducational schools and their long term impact on students' self-esteem, social, and emotional well-being, as well as the effect it has on future career fields and choices of participants. There is very little current research as it relates to this area in the United States today. While a few older international studies exist, the overall results are mixed and not very meaningful to the American education system or culture today. A qualitative study which examines student' attitudes, self-esteem, and emotional well-being in American coeducational schools with single gender classes would be extremely valuable in determining the long-term viability of this relatively new educational initiative. While the current study clearly shows that single gender instruction can be a viable alternative academically, more research is needed to determine the long term impact to students in areas other than instruction.

A final avenue of possible research centers on the investigation of total single gender schools and their overall impact on student academic achievement and academic well-being. While this area of investigation would be very beneficial and unique for researchers, it is not a viable option in American schools today. Federal regulations prohibit mandatory assignment by gender and thus severely limit this option (U.S. Dept. of Education, 2006).

**Limitations**

For this research study, various limitations were pinpointed that could possibly affect the overall validity of the project. Five limitations were recognized for their impact on the internal validity of the study and four limitations were noted concerning the external validity of the research being performed.

First and foremost, the history or setting in which the research took place is very important. Multiple schools, climates, cultures, teacher efficacies, and leadership styles were used in the twelve educational settings utilized for this project. Since this study focused on a large number of subjects and the overall impact that the type and level of instruction has on student achievement, these variables were not examined or controlled for. This type of limitation affects both the type and level of instruction (Gall et al., 2007)

Secondly, the lack of random assignment for students also affects the internal validity of the research project. Due to federal regulations and student opt out provisions the researcher or school districts involved have no way to assign these students using a random method (U.S. Dept. of Education, 2006). This limitation only affects the data

involving type of instruction as students are assigned to the level of instruction based upon test scores and cognitive ability data (Gall et al., 2007).

Additionally, "experimental treatment diffusion" is another possible imperfection which threatens the internal validity of this study (Gall et al., 2007). In all twelve middle schools, single gender instruction was taking place alongside coeducational classes. As teachers meet in professional learning teams and share instructional strategies and ideas, it is plausible that single gender techniques were taken and employed in the coeducational classroom. Again, due to state and federal policies and guidelines, this is an unavoidable drawback of the study (U.S. Dept. of Education, 2006). This limitation only affects the type of instruction being delivered in the schools.

Yet another plausible threat to the overall internal validity of the study is "compensatory rivalry" among students in coeducational or single gender classes (Gall et al., 2007). Since all twelve middle schools had both single gender and coeducational classes going on together, it is conceivable that students from either group could see the other as rivals instructionally and work harder to try and overcome perceived differences in the instructional setting. This limitation affects both the type and level of instruction; however, it is unlikely to be as strong for the level of instruction since students in South Carolina have been separated by level beginning in the third grade.

A final potential inadequacy of this study centers on the possibility of "statistical regression" and the impact on the overall results concerning the level of instruction (Gall et al., 2007). This limitation is possible due to the large differences present in the pre-test scores from the fall of 2011. With any large disparity in initial test scores, it is always

plausible, however unlikely, that the difference measured was due to this statistical likelihood, as opposed to the differences in the overall instruction of the students.

While all five of these internal validity limitations are legitimate concerns for this study, the researcher did not see any of them to be significant factors in the overall results or legitimacy of the research project.

In addition to the internal validity concerns, four limitations were noted as possible factors affecting the overall external validity of the study. The major external validity concern centered on "population validity" and the ability of the researcher to generalize the overall results to a larger cluster of students than merely the 2,079 children examined for this project (Gall et al., 2007). Based on the size of the sample and the ability of the researcher to draw from urban, suburban, and rural settings in the state of South Carolina, the investigator feels comfortable in generalizing this sample to the Southeastern United States as a whole. While the examiner would like to be able to generalize this sample to the United States as a whole; educational values, climates, cultures, and norms vary widely across our country and this study does not enable the researcher to make these generalizations.

A second but unlikely external validity threat involves the possible ecological concern of the "novelty effect" surrounding single gender classes and their implementation in these twelve schools (Gall et al., 2007). While this was a feasible concern, the fact that all twelve schools in the study had implemented single gender classes at least two years prior made this limitation very small and highly unlikely.

An additional limitation and reasonable risk to the external validity centered on the ecological factor of experimenter bias that impacted each and every school

individually and affected the overall student achievement. This limitation was very similar to the internal imperfection regarding the history and setting. With any educational initiative, the overall teacher efficacy and fidelity of implementation were considered major factors in the success of the educational program being studied. The size of the study and the strong instrumentation being utilized were attempts by the researcher to minimize the effects of this limitation (Gall et al., 2007).

A final minor external validity limitation involved the "interaction of time of measurement and treatment effects" for the project (Gall et al., 2007). In this study, the fall 2011 MAP tests and the spring 2012 MAP tests were all not administered on the exact same day in all twelve schools. Due to technological constraints, student holidays, slightly dissimilar testing windows, and differing school district calendars, testing variations were observed. These differences were minute in nature and did not significantly impact the overall number of instructional days each student received. Furthermore, in each individual school setting, coeducational and single gender students received an equal number of instructional days, further mitigating this factor.

**Conclusion**

The results of this large scale research project clearly show that single gender instruction can be a viable alternative to coeducational instruction in eighth grade mathematics classes. While there was not a significant difference in achievement, single gender students did outperform their coeducational peers in both the honors and non-honors setting. Furthermore, the findings of this study also indicate that ability based grouping can have a positive impact on lower academically functioning students and allow them to make strides in lowering the achievement gap between themselves and

their higher functioning peers. In this study of over 2,000 subjects, non-honors' students showed a significant difference in achievement, when compared with their honors' classmates, outperforming their honors' peers in single gender and coeducational classes.

As a result of this study, the researcher has learned and reaffirmed some very meaningful educational insights as well. First and foremost, in order for education to continue to advance and better serve and instruct our students in the twenty-first century, educational research is vitally important and should be supported and promoted by all public and private academic institutions. While the researcher received outstanding support from numerous school districts in the state of South Carolina, many other districts refused to share their large bodies of data on single gender instruction. Educational leaders must promote the advancement of knowledge and not hinder it in any way. Secondly, the researcher also learned the importance of continually re-examining and investigating age old practices such as ability grouping to ascertain their current viability in the twenty-first century educational setting. Learning is a fluid and ever-changing process and instructional strategies and techniques that were ineffective decades ago can be proven successful today in a new and very different educational culture and climate. Finally, as a result of this research project, the investigator plans to present the detailed findings of the study to the superintendent and district office staff to encourage the possibility of expanding single gender classes in middle school mathematics to cover all three grade levels. The researcher is also examining the fiscal viability of creating ability grouped enrichment classes in mathematics for both middle and high school students to promote better mathematics achievement for all students.

In closing, the discoveries from this research add to the current body of data by providing researchers with a large scale, quantitative study concerning the impact that the type and level of instruction have on academic achievement in eighth grade mathematics utilizing nationally recognized instrumentation.

References

Boyd-Zaharias, J., & Pate-Bain, H. (2008). Class matters – In and out of school. *Phi Delta Kappan, 90*(1), 40 – 44.

Brooks, T. (2011). *Effect of single-gender middle school classes on science achievement and attitude* (Doctoral dissertation). Retrieved from ProQuest Dissertations (3445013).

Bruce, C.D., Esmonde, I., Ross, J., Dookie, L., & Beatty, R. (2010). The effects of sustained classroom-embedded teacher professional learning on teacher efficacy and related student achievement. *Teaching and Teacher Education, 26,* 1598 – 1608.

Brutsaert, H. (2006). Gender-role identity and perceived peer group acceptance among early adolescents in Belgian mixed and single-sex schools. *Gender and Education, 18*(6), 635 – 649.

Burris, C.C., Heubert, J.P., & Levin, H.M. (2006). Accelerating mathematics achievement using heterogeneous grouping. *American Educational Research Journal, 43*(1), 105 – 136.

Carroll, A. (2002). At-risk and not at-risk adolescent girls in single-sex and mixed-sex school settings: An examination of their goals and reputations. *Westminster Studies in Education, 25*(2), 147 – 162.

Computer-Based Adaptive Assessments. Retrieved July 24, 2011 from http://www.nwea.org/products-services/computer-based-adaptive-assessments

Elam, Jeanette H. (2009). *An examination of single-gender and coeducational classes: Their impact on the academic achievement of middle school students enrolled in mathematics and science at selected schools in Georgia* (Doctoral dissertation). Retrieved from ProQuest. (3421367.)

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An Introduction* (8th ed.). Boston: Pearson Education Inc.

Gary, S. D. (2011). *A study of the impact on student achievement of single-gender classes in coeducational public schools* (Doctoral dissertation). Retrieved from Proquest Dissertations (3487206).

Gillibrand, E., Robinson, P., Brawn, R., & Osborn, A. (1999). Girls' participation in physics in single sex classes in mixed schools in relation to confidence and achievement. *International Journal of Science Education, 21*(4), 349 – 362.

Gray, C., & Wilson, J. (2006). Teachers' experiences of a single-sex initiative in a co-education school. *Educational Studies, 32*(3), 285 – 298.

Groves, P. (2005). *Single-gender classrooms: Improving achievement of grade 9 males* (Doctoral dissertation). Retrieved from ProQuest. (MR04635.)

Gurian, M. & Stevens, K. (2004). With boys and girls in mind. *Educational Leadership, 62*(3), 21 – 26.

Hauser, C., & Kingsbury, G. G. (2009, April). *Individual score validity in a modest-stakes adaptive educational testing setting.* Paper presented at the annual meeting of the National Council on Measurement in Education, San Diego, CA.

Hoffman, B. H., Badgett, B. A., & Parker, R. P. (2008). The effect of single-sex instruction in a large, urban, at-risk high school. *The Journal of Educational Research, 102*(1), 15 – 35.

Houston, T. (2011). *The impact of single gender education in South Carolina middle schools with grades six through eight: A quantitative study comparing student performance, student attendance, student discipline, and teacher, student, and parent perceptions before and after implementation of single gender education* (Doctoral dissertation). Retrieved from ProQuest. (3460464.)

Howell, D. C. 2008). *Fundamental statistics for the behavioral sciences* (7th ed.). Belmont, CA: Wadsworth, Cengage Learning

Hughes, T. A. (2006). The advantages of single-sex education. *National Forum of Educational Administration and Supervision Journal, 23*(2), 5 – 14.

Jackson, C. (2002). Can single-sex classes in co-educational schools enhance the learning experiences of girls and/or boys? An exploration of pupils' perceptions. *British Educational Research Journal, 28*(1), 37 – 48.

James, A. N. (2009). *Teaching the female brain: How girls learn math and science.* Thousand Oaks, CA: Corwin Press, Inc.

James, A. N. (2007). *Teaching the male brain: How boys think, feel, and learn in school.* Thousand Oaks, CA: Corwin Press, Inc.

Kessells, U. & Hannover, B. (2008). When being a girl matters less: Accessibility of gender-related self-knowledge in single-sex and coeducational classes and its impact on students' physics related self-concept of ability. *British Journal of Educational Psychology, 78,* 273 – 289.

Malacova, E. (2007). Effect of single-sex education on progress in GCSE. *Oxford Review of Education, 33*(2), 233 – 259.

Martino, W., Mills, M., & Lingard, B. (2005). Interrogating single-sex classes as a strategy for addressing boys' educational and social needs. *Oxford Review of Education, 31*(2), 237 – 254.

Mertler, C.A., & Vannatta, R.A. (2005). *Advanced and multivariate statistical methods* (3rd ed.). Glendale, CA: Pyrczak Publishing.

Mojavezi, A. & Tamiz, M. P. (2012). The impact of teacher self-efficacy on the students' motivation and achievement. *Theory and Practice in Language Studies, 2*(3), 483 – 491.

Mulholland, J., Hansen, P., & Kaminski, E. (2004). Do single-gender classrooms in coeducational settings address boys' underachievement? An Australian study. *Educational Studies, 30*(1), 19 – 32.

Nelson, A. L. (2012). *The relationship between middle school teachers' perceptions of principals' transformational leadership practices, teachers' sense of efficacy and student achievement* (Doctoral dissertation). Retrieved from Proquest Dissertations (3530736)

Nielsen, D. C., Barry, A. L., & Staab, P. T. (2008). Teachers' reflections of professional change during a literacy-reform initiative. *Teaching and Teacher Education, 24*(5), 1288 – 1303.

NWEA 2011 Norms Study Data, Retrieved March 30, 2013 from http://www.nwea.org/support/article/rit-scale-norms-study-data-files

Parker, L. H. , & Rennie, L. J. (2002). Teachers' implementation of gender-inclusive instructional strategies in single-sex and mixed-sex science classrooms. *International Journal of Science Education, 24*(9), 881 – 897.

Reliability of NWEA Results, Retrieved July 25, 2011 from http://www.nwea.org/support/article/533/reliability-nwea-results

Reliability and Validity Estimates: NWEA Achievement Level Tests and Measures of Academic Progress, (2004, March). Retrieved July 21, 2011 from http://images.pcmac.org/Uploads/Jacksonville117/Jacksonville117/Sites/DocumentsCategories/Documents/Reliability_and_Validity_Estimates.pdf

Sample South Carolina Eighth Grade Mathematics MAP Questions, (2011). Pearson Education Systems, Retrieved from https://georgetown.schoolnet.com

Saunders, R. (2005). *A comparison study of the academic effects of ability grouping versus heterogeneous grouping in mathematics instruction* (Doctoral dissertation). Retrieved from Proquest Dissertations (3194965)

Sax, L. (2005). *Why Gender Matters.* New York: Broadway Books.

Schneeweis, N., & Zweimuller, M. (2012). Girls, girls, girls: Gender composition and female school choice. *Economics of Education Review, 31*, 482 – 500.

Smith, D. M. (2010). *A comparison of single-gender classes and traditional, coeducational classes on student academic achievement, discipline referrals, and attitudes towards subjects* (Doctoral dissertation)**.** Retrieved from ProQuest Dissertations (3413122)

Solomon, C.B. (2007). *The relationships among middle level leadership, teacher commitment, teacher collective efficacy, and student achievement* (Doctoral dissertation). Retrieved from ProQuest Dissertations (3322742)

South Carolina State Department of Education (2010). *South Carolina middle-school report cards.* Retrieved from http://ed.sc.gov/topics/researchandstats/schoolreportcard/

Spielhagen, F.R. (2010). Algebra for everyone? Student perceptions of tracking in mathematics. *Middle Grades Research Journal, 5.4,* 213 – 223.

Spielhofer, T., Benton, T., & Schagen, S. (2004). A study of the effects of school size and single-sex education in English schools. *Research Papers in Education, 19*(2), 133 – 159.

Spikes, E. S. (2008). *The effects of single-gender classrooms and previous middle school gifted math experience on the mathematics achievement of gifted girls in public education* (Doctoral dissertation). Retrieved from Proquest Dissertations (3325350)

Sullivan, A. (2009). Academic self-concept, gender, and single-sex schooling. *British Educational Research Journal, 35*(2), 259 – 288.

Tieso, C.L. (2000). *The effects of grouping and curricular practices on intermediate students' math achievement* (Doctoral dissertation). Retrieved from Proquest Dissertations (9982009)

Tschannen-Moran, M. & Barr, M. (2004). Fostering student learning: The relationship of collective teacher efficacy and student achievement. *Leadership and Policy in School, 3*(3), 189 – 209.

Turnbull, B. (2002). Teacher participation and buy-in: Implications for school reform initiatives. *Learning Environments Research, 5*(3), 235 – 252.

U.S. Department of Education. (1983). *A nation at risk.* Retrieved from http://www2.ed.gov/pubs/NatAtRisk/index.html

U.S. Department of Education. (2001). *No child left behind.* Retrieved from http://www2.ed.gov/nclb/landing.jhtml

U. S. Department of Education. (2006). *Nondiscrimination on the basis of sex in education programs or activities receiving federal financial assistance; final rule/34, CFR Part 106.* Retrieved from http://www2.ed.gov/legislation/Fedregister/finrule/2006-4/102506a.html

U.S. Department of Justice (1972). *Title IX of the education amendments of 1972.* Retrieved from http://www.justice.gov/crt/about/cor/coord/titleixstat.php

Van de Gaer, E., Pustjens, H., Van Damme, J., & De Munter, A. (2004). Effects of single-sex versus co-educational classes and schools on gender differences in progress in language and mathematics achievement. *British Journal of Sociology of Education, 25*(3), 307 – 322.

Van Zyl, D.G. (2011). *A study of single-gender grouping for sixth grade math as a strategy for improving student achievement* (Doctoral dissertation). Retrieved from Proquest Dissertations (3515846)

Vrooman, M. K. (2009). *An examination of the effects of single-gender classes on reading and mathematics achievement test scores of middle school students* (Doctoral dissertation). Retrieved from Proquest Dissertations (3371729)

Warrington, M. & Younger, M. (2003). 'We decided to give it a twirl': Single-sex teaching in English comprehensive schools. *Gender and Education, 15*(4), 339 – 350.

Warrington, M. & Younger, M. (2001). Single-sex classes and equal opportunities for girls and boys: Perspectives through time from a mixed comprehensive school in England. *Oxford Review of Education, 27*(3), 339 – 356.

Whalen, W.V. (2012). *The single-gender classroom: Improving middle school students' achievement in math* (Doctoral dissertation). Retrieved from Proquest Dissertations (3525858)

Wong, K. C., Lam, Y. R. & Ho, L. M. (2002). The effects of schooling on gender differences. *British Educational Research Journal, 28*(6), 827 – 843.

# APPENDIX

# APPENDIX A

## Sample South Carolina Eighth Grade Mathematics MAP Questions

**Removed for Publication**

# APPENDIX B

## Institutional Review Board Approval Letter

## Institutional Review Board Approval Letter

LIBERTY UNIVERSITY

The Graduate School at Liberty University

October 31, 2012

David Hammel
IRB Exemption 1442.103112: The Effect of Single Gender Instruction on Eighth Grade Students' Mathematics Achievement

Dear David,

The Liberty University Institutional Review Board has reviewed your application in accordance with the Office for Human Research Protections (OHRP) and Food and Drug Administration (FDA) regulations and finds your study to be exempt from further IRB review. This means you may begin your research with the data safeguarding methods mentioned in your approved application, and that no further IRB oversight is required.

Your study falls under exemption category 46.101 (b)(4), which identifies specific situations in which human participants research is exempt from the policy set forth in 45 CFR 46:

> (4) Research involving the **collection or study of existing data**, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or **if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects**.

Please note that this exemption only applies to your current research application, and that any changes to your protocol must be reported to the Liberty IRB for verification of continued exemption status. You may report these changes by submitting a change in protocol form or a new application to the IRB and referencing the above IRB Exemption number.

If you have any questions about this exemption, or need assistance in determining whether possible changes to your protocol would change your exemption status, please email us at irb@liberty.edu.

Sincerely,

**Fernando Garzon, Psy.D.**
*Professor, IRB Chair*
**Counseling**

**(434) 592-4054**

LIBERTY UNIVERSITY

*Liberty University | Training Champions for Christ since 1971*