12-2014

Ability Grouping in Elementary Education

Anna Kathryn McCarter
East Tennessee State University

Ability Grouping in Elementary Schools

A dissertation

Presented to

The faculty of the Department of Educational Leadership and Policy Analysis

East Tennessee State University

In partial fulfillment

Of the requirement for the degree

Doctor of Education of Educational Leadership

by

Anna Kathryn McCarter

December 2014

Dr. Virginia Foley, Chair

Dr. Cecil Blankenship

Dr. Eric Glover

Dr. James Lampley

Keywords: ability grouping, homogeneous, heterogeneous, academic achievement

ABSTRACT

Ability Grouping in Elementary Education

by

Anna Kathryn McCarter

Ability grouping in education is a common practice used to differentiate instruction in order to meet the academic needs of students. The primary purpose for grouping students by ability is to increase their academic growth and achievement by providing instruction at the students' current instructional level. However, there is much conflicting research regarding the impact of grouping students by ability and its link to student achievement. The purpose of this study was to determine if there is a statistically significant difference between school-wide student achievement in grades 3, 4, and 5 based on the type of grouping: ability grouped or not grouped by ability classroom settings. Data were collected from the Tennessee Department of Education website for the 2012-2013 academic school year as well as from individual school administrators regarding how students were grouped for instruction: grouped by ability or not grouped by ability. Independent samples t-tests were run to determine if there is a significant difference between students who received instruction in ability grouped setting and those that were not placed in an ability grouped classroom. The results of this study indicated that there is no difference in achievement scores based on the type of instructional setting (ability grouped or not grouped by ability) in reading and math in grades 3, 4, and 5.

DEDICATION

This study is dedicated to my family, friends, and coworkers who continually encourage and challenge me to be the best individual, wife, mother, daughter, friend, and teacher that I can be.

To my husband Bryan, you always encourage me to grow and push me to be my best in everything that I do. You have supported me, encouraged me, listened to my frustrations, and taken care of our family so that I could focus my energy on this project. I am grateful for your constant positivity, laughter, and support.

When I began this journey I was not a mother, but 3 years ago I had the blessing of welcoming my daughter Mattie Kathryn into this world. You are a light in my life and you bring joy every day. Thank you for being patient while I have worked through this program and for always giving me a hug and smile when I need it most.

As a young child my parents, Warren and Marcy White, always encouraged me to work hard, pursue my dreams, and do something that I love. I am grateful for the example that they have provided. Thank you for supporting me as I have walked through this journey. Your prayers, words of encouragement, and belief in me have had a pivotal role in developing me into the woman that I am today, and for that I am grateful.

ACKNOWLEDGEMENTS

I would like to acknowledge all of my family, friends, and coworkers. I am blessed to be part of your life and am grateful for your support and encouragement as I have progressed through this journey.

To the department of Educational Leadership and Policy Analysis faculty and staff: I am thankful for the opportunity to learn and grow under such wise leadership. Thank you for pressing me to think outside the box and continually grow as an educator. Your encouragement and guidance through this program have been pivotal in my career as an educator.

I am especially grateful to my committee members: Dr. Cecil Blankenship, Dr. Eric Glover, and Dr. James Lampley. Thank you for your support, encouragement, and guidance as I have progressed through this process. Your time and wisdom throughout this process are greatly appreciated. To my chair Dr. Foley: thank you for pushing me to complete this program and for believing in me. Your patience, encouragement, and advice have caused me to grow throughout this program, and for that I am grateful.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

It is estimated that in the year 2015 English Language Learners will comprise approximately 30% of the public school population, therefore increasing the need for differentiated instruction. Christie (2008) found herself in a conundrum a few years ago when her son befriended a new student from Vietnam. During their first year as friends the two boys were paired together in a classroom not grouped by ability during math because they often excelled (Christie, 2008). The boys became friends and spent much time together; however, when the next school year came around the young student from Vietnam was not placed in the advanced math class with his friend. When approached, the principal responded that the lack of English skills would prevent this student from excelling in the advanced math class. Unfortunately, the two boys became distant and did not spend much time together anymore because they were no longer in the same academic classes. In high school the young Vietnamese student was placed in "baby" classes and on the vocational track, whereas his friend was placed in the university track. Placing students into classes based on their ability creates friendships (Flashman, 2012), which can have an impact on their academic achievement.

The practice of placing students into small groups in order to meet their academic needs is a common practice in education. For centuries educators have placed students into groups based on their strengths, weaknesses, or areas of need (Ansalone, 2012; Betts, 2011; Catsambis & Buttaro, Duflo, Dupas, & Kremer, 2009; 2012; Flashman, 2012). In some cultures students are placed into classroom not grouped by ability, and teachers have small group instruction based on student need. Other cultures place students into ability-grouped classes based on academic ability or achievement (Emery, 2007). There has been much research supporting both ability-

grouped classrooms and classroom not grouped by ability (Catsambis & Buttaro, 2013; Duflo, Dupas, & Kremer, 2009; Flashman, 2012). Ability grouping has been defined as placing students in classrooms based on their ability in a particular subject, and classroom not grouped by ability has been defined as placing students in mixed-ability classrooms. Regardless of the type of ability grouping implemented, it is clear that the primary purpose of grouping students into small groups based on their academic ability is to increase student learning as well as to make differentiation simpler for teachers.

Based on the current data-driven educational system educators are constantly making changes in their classrooms in order to meet the ever-increasing demands to raise student achievement as measured by standardized test scores (Lee, 2008; Musoleno, 2010; Popham, 2007, 2011; Slavin, 1987). During President George W. Bush's administration No Child Left Behind legislation was signed into law and implemented. Part of the legislation required a focus on closing the achievement gap in each school based on its demographics and achievement scores. This reauthorization of ESEA added pressure for educators to increase student achievement and growth, therefore causing educators implement new instructional strategies. Ability grouping, both within classrooms and between classrooms, showed positive and negative impacts on students academically, socially, and emotionally (Reuman, 1989). While Duflo et al. (2009) showed that there are academic benefits to grouping students into ability grouped classrooms, other research indicated that there is potential to see a decrease in self-esteem in students who experience this type of ability grouping (Reuman, 1989). Overall, educators are not only responsible for providing a strong academic experience but also for developing successful members of society and caring for each individual student. Therefore, when making decision

regarding instructional strategies educators must evaluate the positive and negative impact of their choice.

Statement of Problem

With the ever-increasing changes in legislation educators are using many different instructional practices in order to meet the high-demands placed upon them (Hornby, Witte, & Mitchell, 2011). The practice of ability grouping has historically been used to group students and place them into tracks according to their academic abilities and their career paths. However, with the changes in high-stakes testing the motivation has morphed from placing students into classes that will prepare them adequately for their future career to using standardized data to place students in a class in which they grow academically according to another standardized test (Emery, 2007). With conflicting research regarding the academic benefits of ability grouping (Duflo et al., 2008; Reuman, 1989), educators have looked to the social impact of ability grouping. Many researchers found that there are negative impacts associated with ability grouping and student self-concept (Boeler, 2008; Catsambis, 2011; Fendler & Muzaffar, 2008; Kintz, 2011). Further, researchers discovered that when students are placed into particular ability groups, social comparison occurs and that many times students are "tracked" for a particular career path. In order to meet the demand placed on schools and educators to increase student achievement and close the achievement gap many schools have implemented instructional strategies with the purpose of meeting students' academic needs based on standardized test scores. One of the strategies that educators have used to meet these targeted needs and to increase achievement based on standardized assessments is through the practice of ability grouping. Placing students into groups or classrooms based on achievement is a common practice worldwide. Students in the United Kingdom are placed into tracks upon starting school.

Most Asian countries place students into schools based on their performance on a standardized assessment. In the United States there are differing approaches to ability grouping (Boeler, 2008; Catsambis, 2011; Fendler & Muzaffar, 2008; Kintz, 2011). In secondary schools students tend to follow either a vocational or university track. However, limited research has been conducted on the impact that ability grouping has on elementary academic achievement.

Studies conducted regarding elementary ability grouping showed that the practice of ability grouping does not increase academic achievement in reading or in math (Lleras & Rangel, 2009; Reuman, 1989). According to Lleras and Rangel students placed in low achieving reading groups actually learn less every year as they progress through school, thus increasing the achievement gap. Research regarding mathematics revealed that there is not a significant difference in student achievement solely based on ability grouping; however, it is much more common for teachers to ability group between classes for math.

This study was an evaluation of student achievement on the Tennessee Comprehensive Assessment Program (TCAP) in reading and math for grades 3, 4, and 5 comparing students who received instruction in ability grouped classrooms to those who received instruction in a classroom not grouped by ability. The purpose of this study was to determine if there is a statistically significant difference between school-wide student achievement in grades 3, 4, and 5 based on the type of grouping: ability grouped classrooms or classrooms not grouped by ability.

Research Questions

This study was guided by the following research questions:

Research Question 1

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic,

Below Basic) for third grade math between schools that have students grouped by ability and schools that have students not grouped by ability?

Research Question 2

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for fourth grade math between schools that have students grouped by ability and schools that have students not grouped by ability?

Research Question 3

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for fifth grade math between schools that have students grouped by ability and schools that have students not grouped by ability?

Research Question 4

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for third grade reading between schools that have students grouped by ability and schools that have students not grouped by ability?

Research Question 5

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for fourth grade reading between schools that have students grouped by ability and schools that have students not grouped by ability?

Research Question 6

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for fifth grade reading between schools that have students grouped by ability and schools that have students not grouped by ability?

Significance of the Study

Administrators and teachers are consistently faced with making decisions regarding instructional practices that impact student learning. One of the commonly used practices is grouping students into classrooms based on their ability for instruction in both reading and math. In order to make decisions that reflect research-based practices it is important to determine if there is a significant difference in the percentage of students in each TCAP category (Advanced, Proficient, Basic, Below Basic) between schools that place students into classrooms for instruction based on their ability and those that place students into classrooms with students of varied ability. This study provided an evaluation of the difference in TCAP scores based on the type of grouping implemented in schools; therefore, it provided more information for educators to use in making instructional decisions.

Definitions of Terms

In education there are many terms that are used in association with education practices. Below are the operational definitions for terms important to this study.

Ability Grouping: students are placed into classes based on their past achievement and ability to learn (also known as homogeneous grouping)(ASCD, 2014).

Achievement Gap: Lasting differences in achievement amongst different subgroups of students as shown by standardized test scores, class grades, and other data (ASCD, 2014).

Achievement Tests: Tests used to calculate student learning in different school subjects. "These norm-referenced, multiple-choice tests are intended to measure students' achievement in the basic subjects found in most school districts' curriculum and textbooks. Results are used to compare the scores of individual students and schools with others—those in the area, across the state, and throughout the United States" (ASCD, 2014, para. 8).

Banding: The practice of grouping students into classes based on ability; however, students can be regrouped for some subjects. For example, students may be in band 1 for reading but band 2 for math (Wouters et al., 2012).

Heterogeneous Grouping: Placing students of varying talents and academic ability in the same classroom for instruction (the opposite of homogeneous grouping) (ASCD, 2014).

Homogeneous Grouping: Placing students in classes based to their abilities and academic ability (ASCD, 2014).

Setting: The practice of grouping students between classes based on general ability for each subject. This is the most flexible form of between class ability grouping and it is

possible for a student to be in a different group for all subject areas (Wouters et al., 2012).

Streaming: The practice of grouping students between classes based on ability. Students are usually in the same group for all subject areas (Wouters et al., 2012).

Tracking: "The practice of dividing students for instruction according to their perceived abilities. Students are placed on a particular track (college-bound, general, vocational, and remedial) and given a curriculum that varies according to their perceived abilities and future positions in life. At the elementary level, the practice is called grouping." (ASCD, 2014, para. 18).

Delimitations and Limitations

1. The data were limited to school achievement scores found on the Tennessee Department of Education website of school report cards.
2. There classification of grade level grouping was limited to the principal's response when contacted regarding ability grouping at each school.
3. Schools were classified based on phone calls placed to school administrators and schools were coded accordingly.
4. Data were limited to the 2012-2013 year as this was the most current data available at the time of analysis.
5. Accuracy of coding the type of ability grouping is based on principals understanding grouping definitions and accurately reporting the type of grouping.

Overview of the Study

A quantitative approach was taken to evaluate the relationship between grouping practices and academic achievement in grades 3, 4, and 5 in both reading and math. Data were collected from the Tennessee Department of Education Website. Percentages of students scoring in the categories of advanced, proficient, basic, and below basic from the school report cards for the 2012- 2013 school year were collected in the subject areas of reading and math. The school report card provides percentages of schools that score in each proficiency level in each subject on the Tennessee Comprehensive Academic Program (TCAP). Also, principals were contacted to determine if the school implemented ability grouping or did not group by ability in grades 3, 4, and 5 in reading and math. The schools were limited to East Tennessee; all schools were located east of Knoxville. The sample of schools was selected based on the criteria that they grouped their students into ability grouped classrooms or classrooms not grouped by ability as a grade level for both reading and math.

CHAPTER 2

REVIEW OF LITERATURE

Ability grouping is the practice of placing students into groups based on their academic achievement level (Slavin, 1990). Historically, teachers have used this instructional practice in secondary schools and placed students into tracks such as vocational or university according to their career choice and academic capabilities. As high-stakes testing has become more demanding and influential in education, educators have implemented ability grouping as an avenue to increase academic achievement. Researchers have found conflicting evidence regarding the overall impact of ability grouping; however, most researchers concluded that ability grouping does not have a positive impact on student achievement for students in reading and had very little effect on students in math based on standardized test scores. Nevertheless, researchers found that ability grouping is implemented in order to differentiate instruction toward a specific level of students and is often initiated based on personal beliefs rather than academic research (Hornby et al., 2011).

Ability grouping is a common practice that is typically used to differentiate instruction with the goal of increasing student achievement (Hornby et al., 2011). However, despite the numerous research studies conducted on this topic it is clear that there is conflicting research regarding the impact that ability grouping has on student achievement as well as student self-concept.

Ability grouping has been used by many classroom teachers in order to meet the needs of students and increase overall academic growth. According to Chorzempa and Graham (2006) ability grouping is the instructional practice of breaking into small groups based academic achievement level. However, there is research that has indicated that the type of ability grouping,

either between classrooms or within a classroom, makes a difference in student growth throughout the school year (Reuman, 1989). Further, many researchers have stated that ability grouping is a social injustice and decreases the self-concept that students have of themselves and of their academics (Fram, Miller-Cribbs, & Van Horn, 2007). Research has been conducted to determine if there is social inequity when implementing the practice of ability grouping. Catsambis, Mulkey, Buttaro, and Steelman (2011) conducted a study in which teacher placed students into reading groups in Kindergarten. They found that boys and girls are equally placed into average groups; however, boys were underrepresented in high achieving groups and girls were underrepresented in low achieving groups. Kulik (1992) conducted a meta-analysis of literature regarding ability grouping and found that the amount of growth that occurs is determined by the following: all classrooms following the same program, programs that adjust curriculum to the specific ability, and programs that make specific adjustments to the curriculum for the special needs and gifted and talented students. In essence, when grouping students, either within the classroom or between classrooms, Kulik found that the amount of alteration made to the curriculum impacts the academic achievement. The highly debated topic of ability grouping has brought forth research that both supports and refutes ability grouped class settings and class settings not grouped by ability.

Definition of Ability Grouping

The concept of ability grouping has been used in education worldwide. However, there are many differences based on culture and varying vocabulary that surrounds this topic. According to Wouters et al. (2012) the practice of ability grouping has been identified as a form of differentiating instruction in order to meet the academic needs of students. The British refer to

ability grouping as streaming or tracking, which is a form of between-class ability grouping based on general ability. Banding is similar to streaming, yet it is more flexible and students are regrouped based on the particular subject area. Setting, also known as regrouping, is the least restrictive, and students are grouped based on their ability in each respective content area (Ireson & Hallam, 1999, Wouters et al., 2012).

Many schools determine the path that students will take academically based on their performance on a standardized assessment, and students are grouped by strengths into ability-grouped classrooms. Schofield (2010) found that students are not always categorized strictly based on their intelligence but also by socioeconomic status, race, and behavior. Between-class ability grouping is typically seen in secondary settings; however, it is starting to appear in the elementary level in the United States. Conversely, the concept of within-class grouping is often used at the primary level in order differentiate instruction and meet the needs of the heterogeneous classroom. The heterogeneous classroom is a traditional classroom with mixed abilities and ability grouping takes place in small groups.

Schofield (2010) highlighted a significant difference between ability grouping and curriculum differentiation. Whereas ability grouping is the practice of placing students into groups, tracks, or classrooms based on their intelligence or ability, curriculum differentiation is defined as adjusting the curriculum to meet the needs of the student. This could be anything from placing students into small groups to altering the curriculum entirely (Schofield, 2010).

Historical Context for Ability Grouping

Historical Roots

Placing people into groups based on ability or interest is not a new topic in education or in the workforce. There has been much debate on the most fruitful composition of grouping: heterogeneous or homogeneous. Placing students into tracks in schools became popular during the Progressive period (1890-1920) when industrialization was increasing. The local school boards encouraged schools to place students into either a vocational track or a middle management track (Emery, 2007). The need for a strong working class along with the influx of immigrants in need of learning English led educators to place students into groups based on their career paths and to empower immigrants both socially and politically. Ansalone (2010) asserted ability grouping was used at the turn of the century in order to "Americanize" the southern African Americans and influx of immigrants.

After taking more control through several legislative acts school superintendents were determined to sort and socialize students to meet the demands of the industrialized era (Emery, 2007). In order to do this standardized intelligence testing was created. Historically there have been many shifts in education with changes toward a national curriculum as well as assessment. Ireson and Hallam (1999) attributed the need for ability grouping to this push for more accountability and standardized testing. In the 1990s ability grouping became much more popular, contrary to the previous 2 decades when heterogeneous classrooms were recommended. Grouping students into classrooms based on their ability has been implemented as a response to the increase in accountability and drive to meet the competitive global community (Ireson & Hallam, 1999).

Motivation for Ability Grouping

Typically the type of ability grouping is a decision that is left to the school administrator. Turney found that there are several commonly used motivations for implementing ability grouping: planning individualized instruction, teacher modification of lessons to class ability level, reduction of the amount of boredom in advanced students, and encouragement of struggling students to participate (as cited in Slavin, 1990).

There are many other motivations for grouping students; primarily, administrators believe that it will help teachers to individualize instruction and increase student achievement (Hornby et al., 2011). Hornby et al. and Forgasz (2010) found that one of the primary motivations for implementing ability grouping into a school setting is to make it easier for teachers to teach a group of students at the same level of instruction. However, Hornby et al. discovered through in-depth interviews with school administrators that more often than not administrators made the decision to ability group based on their personal belief or feelings toward the topic. Forgasz approached the subject matter with a survey of Australian teachers. Forgasz found that while 80% of teachers implemented some form of ability grouping almost 30% of teachers did not agree with the practice. These two researchers have indicated that while administrators are making decisions regarding ability grouping (Hornby et al., 2011) there are approximately 30% of teachers who do not believe in the practice.

After conducting a meta-analysis of literature regarding within-class ability grouping, Lou et al. (1996) determined that based on previous research within class ability grouping is effective and increasing academic achievement. Further, they concluded that as a whole the results indicated that homogeneous grouping proved to have the greatest impact. However, Lou

et al. stated that in order for homogeneous small groups to be effective, the adaptation of instruction and materials is imperative.

Hallam and Ireson (2008) discovered that the teachers' attitude toward ability grouping largely depends on the subject matter they teach. After analyzing survey responses from teachers in over 45 secondary schools, Hallam and Ireson found that heterogeneous classrooms were accepted by teachers in the following subject areas: physical education (86%), design and business (83%), humanities (77%) and English (80%). However, only 49 % of science teachers and 18% of math teachers felt that they could appropriately teach their subject matter to a mixed-ability grouped classroom. The researchers concluded that overall leveled groups are found to be more effective in the eyes of the teacher in math and science and that English, humanities, P.E., and design and business classes do not feel as much of a need for homogeneously grouped classrooms.

Another common use for ability grouping is placing students into small groups for instruction on their particular level. While in reading groups this appeared to be beneficial for students on or above grade level, it was found to have a negative impact on students who were either English Language Learners or performing below grade level (Lleras & Rangel, 2009). Conversely, when ability grouping was implemented into mathematics classrooms, the academic growth was positive (Soloman, 2011).

Teacher Impact on Student Achievement

Further research showed that teachers are the most prominent factor in determining school effectiveness. A recent survey of 1,500 teachers by Calik, Seagin, Kavgaci, and Cagatay-Kilinc (2012) linked teacher's self-efficacy with student achievement as well as a teacher's self-efficacy with principal leadership. Based on this research, it is clear that there is a direct impact

on student achievement, school effectiveness, and teacher self-efficacy that is derived from the principal being a strong leader of learning. As a result of this research Calik et al. (2012) recommended that a leader must know the characteristics of a dynamic instructional leader as well as understand the impact that his or her instructional leadership has on teacher self-efficacy, school culture, and overall school effectiveness.

While relationships and honesty play a large role in increasing teacher self-efficacy, there is much to be said for human values and responsibility. Aslanargun (2012) stated that the principal as an instructional leader has the responsibility to oversee activities regarding the learning process, the teaching-learning process, and student curriculum. After conducting a phenomenological study regarding the values of administrators and how that impacts teachers and students, the top two values identified were justice and confidence. Aslanargun defined justice as being fair and impartial; principals should strive for equality. With regards to teacher self-efficacy, teachers will be more willing to work hard and follow the principal's leadership if they feel that the overall structure is fair. Further, Aslanargun described confidence as being confident in oneself but also having confidence in students and staff. Loyalty was also coupled into this definition as it pertains to the school framework and relationships.

Therefore, in order to see an increase in student achievement and school culture, teacher self-efficacy must be a priority. Houchens (2009) conducted research in which he discovered traits of effective educational leadership. Through this research Houchens found that the greatest component of building an effective school is through trustworthy relationships with teachers and colleagues. He quoted Spillane and Thompson, "Trust was crucial because it facilitated conversations about instructional reform…Moreover, trust created an environment in which educators were comfortable discussion their understandings and reservations about new

instructional approaches…" (as cited in Houchens, 1997, p. 195). The time and effort that it takes to build relationships built on trust is not time spent in vain.

Many studies conducted regarding homogeneous and heterogeneous grouping have studied student growth based on groupings with regards to academic ability; however, Mokhtar, Majid and Foo (2009) conducted a study regarding homogeneous and heterogeneous grouping based on dominate learning styles. Mokhtar et al. discovered that when students were placed in homogeneous and heterogeneous groups based on learning styles, after the students had undergone the learning style training, they performed at a much higher level than those who had not undergone the training. Therefore, while there is much debate regarding the impact of grouping students by ability, it is clear that when students understand how they learn best and implement strategies in independent and group work, academic growth increases.

Regardless of personal beliefs it is clear that administrators must understand the beliefs of their staff and take them into consideration when implementing new initiatives. Further, teachers may not always agree with the decision of the administrator; however, when they have trusting relationships they are more likely to work hard and succeed (Houchens, 2009)

Academic Impact of Ability Grouping

Positive Academic Impact

Perhaps the most widely purported reason for implementing ability grouping is to provide instruction for all students that is affective and meets their needs (Chorzempa & Graham, 2006). In a study regarding the use of within-class ability grouping in reading, Chorzempa and Graham discovered that of the 222 elementary reading teachers surveyed, 63% reported using within-class ability grouping and of that 63%, 68% cited that they used this practice to meet the

students' instructional needs. This research indicated that there is almost two thirds of the population of primary educators who are using within-class ability grouping to aid in the academic achievement of their students.

Fram et al. (2007) found that in regards to first grade reading in rural South Carolina schools, a classroom with a higher population of below-grade level readers actually lowered gains. They found through their study of first graders that those students who are academically able benefit from ability grouping, while those who have lower ability academically lose out. Their reasoning for this is evidenced in the fact that the students with lower ability have lower self-concept and have nothing to work for in order to motivate them in their learning.

In order to turn around a failing school in England, Stewart (2013) studied streaming and decided to incorporate banding into the new educational practices. Students were given a standardized academic achievement assessment that determined their placement in classes. In banding students may be placed in differing classes based on their strengths and weaknesses, whereas in streaming students are placed in one cohort for all subject areas. Stewart found that the school became highly competitive with other schools in England and this program was implemented into the Grammar School Pathway, preparing students for university. Stewart found that by banding students into academic levels the teachers were able to teach to their strengths. Stewart claimed that classrooms that are created with mixed ability are not successful unless there is mixed-ability teaching or differentiation. Stewart further stated that in order for banding to work the weak teachers must not always be placed with the struggling students. There must be teachers who are able to differentiate the instruction and meet the academic needs of students. Regarding his belief in banding, Stewart stated, "The truth is we naturally tend to gravitate towards people of a similar intelligence. Secondly (I said), 'Just look out in the playground, when

students are no longer forced to be with certain people, and see who they choose to congregate with'" (p. 3).

Negative Academic Impact

Slavin (1990) conducted a meta-analysis of 29 international studies regarding ability grouping in secondary schools. All of the researchers collected data that reflected students placement into homogeneous or heterogeneous classrooms by ability as well as students being assessed using a standardized assessment. When converging the studies, Slavin found that there was no correlation between ability grouping and academic achievement. Further, Slavin discovered that contrary to previous research, there was no effect on students of varying ability groups. Slavin reversed the theory that ability grouping for high and low- ability learners leads to increased learning. According to Slavin there is no impact on student achievement when ability grouping is implemented into the school culture.

Kulik (1987) also concluded that while students who were gifted and talented did benefit from homogeneous grouping, all other students did not have statistically significant growth based on their type of grouping. While Kulik did not indicate that there was a negative impact on students, there was not a statistically significant effect regarding their academic achievement as a result of their homogeneous grouping.

Ansalone (2010) completed a review of literature in which he determined that there is no supporting evidence that ability grouping increases academic achievement. While some early researchers indicated that ability grouping increased achievement for high and middle achievers and decreased academic achievement for low performers, the data were not consistent nor were they replicated. Therefore, Ansalone concluded that ability grouping does not increase academic achievement.

Bui, Imberman, and Craig (2012) discovered that students placed in a gifted and talented program did not score significantly higher on a standardized assessment than gifted students not in a gifted and talented program. The researchers conducted two studies. On study was an evaluation of test scores of students enrolled in a gifted and talented program and the other looked at scores of students who barely missed the cut score to enroll, therefore not participating in the gifted and talented program. They found that students enrolled in the gifted and talented program, receiving instruction with other gifted and talented students, did not score significantly higher on assessments; however, they also discovered that students who were right on the cusp of eligibility who received instruction in a mixed-ability group did not score significantly higher either. Overall, the researchers concluded that there was no difference regarding academic achievement between those receiving instruction in a gifted and talented program and those in a heterogeneous classroom.

Reuman (1989) conducted an in-depth study in which he analyzed data from sixth grade students who were placed in differing types of ability-grouped settings. He differentiated students by coding them as receiving instruction in either between-class or within-class ability grouped classrooms. Reuman's research supported the findings of Wouters et al. (2012) in that both studies have found that there is a statistical difference between students receiving between-class and within-class ability grouping and their overall growth. Reuman discovered that students placed in the within-classroom ability grouping setting scored significantly higher on a math achievement test between the beginning and the end of the year than those in the between-classroom ability group setting (1989).

Slavin (1987) found no supporting evidence placing students in homogenous classrooms entirely. He found that homogenous grouping aids in academic instruction in limited amounts.

Slavin suggested that homogenous grouping be used sparingly and most of the school day be spent in heterogeneous classroom settings.

However, the use of ability grouping does allow the instructor to meet the needs of each student where they are and provide instruction at an appropriate pace, therefore increasing growth (Duflo et al., 2009). A study conducted in Kenya by researchers Duflo et al. regarding elementary students showed that students at all levels have experienced academic gains when placed in ability-grouped classrooms. This study discovered that when class size is reduced in conjunction with ability grouping students performing below grade level, on grade level, and above grade level showed academic growth.

There is a large emphasis placed on ability grouping in many educational settings because teachers have been able to tailor their lessons to the ability of the students. Hung, Smith, Harris, and Lockard (2007) discovered that students who have had a history of behavioral problems and have presented distraction in the classroom do not receive the reaction that they have sought from their peers or teachers when they have a specific behavior program and receive behavior modification using a specific technology program. Placing students into ability groups resulted in decreasing the number of behavior disruptions and allowed the teacher to focus on the academic needs of the students, which increases achievement.

As many other researchers found Hornby et al. (2011) discovered that one of the primary motivations of implementing homogenous classroom settings is personal belief. After conducting nine in-depth interviews with administrators in New Zealand Hornby et al. found that administrators were making decisions regarding the practice of ability grouping with limited knowledge regarding research. During the interviews principals shared that ability grouped classrooms seemed to help high achievers because they needed to bounce ideas off of other high

achievers, whereas low achievers struggled because they had no example for higher order thinking and performance. Principals reported that homogenous classrooms also helped students with lower levels of academic achievement because they were in smaller classes with more instructional assistants who were able to tailor instruction to their needs. However, there were no benefits for the average performing student. Hornby et al. suggested that while one of the primary goals of education is to implement research-based practices, many principals are not reading research regarding ability grouping, but rather making this decision based on their own personal beliefs.

Schofield (2010) completed a review of international literature on the impact of ability grouping with regards to the achievement gap. Schofield found that many times when schools or classrooms implemented ability grouping they also employed curriculum differentiation. When the two best practice strategies are put into place, Schofield found that there was actually an increase in the achievement gap. Students receiving instruction at high-level institutions typically had large amounts of growth and low-performing students who received instruction at their respective school struggled and sometime regressed.

Musoleno and White (2010) found that teachers in middle schools are concerned with several new practices being implemented into middle schools as a result of high-stakes testing. The researchers conducted a survey asking questions about instructional practices that were implemented before and after No Child Left Behind. With regards to ability grouping, there were no significant changes in how often students were grouped; however, the survey results revealed that teachers were concerned that student schedules were being modified based on remediation needs. The need for remediation is not only because of a disparity in academic proficiency, but also due to the pressure felt by teachers with regards to testing. The researchers recommended

that in order to maintain the developmentally appropriate structure of middle school, in order to meet the academic, social, and emotional needs of the students, that ability grouping be very flexible and used only throughout portions of the day. This would allow students to receive the tailored instruction, giving teachers time to prepare for high-stakes testing, and also facilitate time for students to be social with students in other ability groups.

Overall, research has shown conflicting results regarding the academic impact of ability grouping. Where elementary students in Kenya have shown growth through the use of ability grouping, Fram et al. (2007) indicated that there has been a decrease in growth for struggling students. Duflo et al. (2008) conducted a study in Kenya regarding ability grouped classrooms and classrooms not grouped by ability. Over 10,000 students were divided into two groups: schools that track based on achievement and nontracked schools. However, this study also included the addition of a teacher as part of the teacher-student ratio reduction program in Kenya. Their class sizes decreased from an average of 86 first graders in a class to 46 students per teacher. After implementing the tracking for 18 months, the test scores were evaluated. Duflo et al. discovered that there was a positive significant difference between students in the tracked schools versus those in the nontracked schools. Conversely, Fram et al. (2007) concluded that while teachers in high-poverty schools spent more time in clustered groups, there test scores were significantly lower than those enrolled in low-poverty schools. Therefore, the study conducted by Fram et al. conflicted with the results of Duflo et al.

Ability Grouping with Regards to High Achievers

A university in Australia implemented specific coursework for preservice teachers regarding gifted and talented learners in order to address the need for differentiated instruction (Plunkett & Kronborg, 2011). Plunkett and Kronborg discovered in their review of literature that

grouping gifted students into homogenous groups helped academically; however, the students need differentiated instruction as well as an opportunity to flourish and grow socially. Therefore, the researchers conducted a study in which they surveyed teachers regarding their perception of teaching gifted and talented learners after they had completed the course at the university. Plunkett and Kronborg discovered that overall the preservice teachers felt that they had a better understanding of gifted and talented students, their social and academic needs, and learned instructional strategies that would benefit the students in a mixed ability classroom. The review of literature found that gifted students needed time with others of like ability but also needed socialization within a mixed ability classroom.

Research has shown many benefits and drawbacks to ability grouping. However, Catsambis and Buttaro (2012) discovered that there might be perks to homogenously grouped classrooms for high achieving students only. Regarding academic placement in Kindergarten, Catsambis and Buttaro found that when placed into small groups for reading, students in high-level groups had a positive disposition regarding reading. However, students placed in low-level groups had a negative disposition toward reading. The researchers analyzed the psycho-social impact that ability grouping had on students and discovered that the high-level achievers had increasingly positive outlooks on education and their abilities that correlated to their academic performance. Conversely, the lower-level achievers had a lower academic performance due to their low outlook on their academic abilities. Similarly Hornby et al., (2011) also found that students in secondary classrooms performed similarly. Students in ability grouped classrooms did not perform significantly higher as a whole with the exception of gifted and talented students or those with special education needs. Overall, it appears that ability groups increase the

motivation and overall success of high-achievers, whereas it negatively impacts low-level achieving students.

Brulles, Saunders, and Cohn (2010) found that students who were homogeneously grouped with other gifted students in grades 2 through 8 saw significant academic gains. The results of this study revealed that as students became older the growth of students placed in these homogeneously grouped classrooms decreased. The researchers categorized gifted students based on their type of instruction: homogenously grouped with other gifted students or heterogeneously grouped. The teachers received professional development for teaching students in a gifted program and differentiation respectively. Overall, the results of Brulles et al.'s study along with the findings of Catsambis and Buttaro (2012) align; there appears to be an association between an increase in academic achievement in gifted learners and homogeneous grouping.

Ability Grouping in Reading and Math

Research has been conducted in order to evaluate if there is a significant difference in student achievement scores when ability grouping is implemented in the subject areas of reading and math. Whereas small group reading based on ability has been historically accepted, Lleras and Rangel (2009) discovered that it is detrimental to struggling students. Students who received instruction in an ability-grouped math classroom showed growth (Soloman, 2011). It is clear that there are differences between the impact of ability grouping in reading and math.

Academic Impact in Reading

Teaching students to read in elementary schools has been one of the primary goals of teachers. As a result teachers in elementary schools have often used ability grouping as a strategy to increase reading success. Lleras and Rangel (2009) found that while ability grouping for

reading has historically been helpful, when low-performing students, especially minority students, are placed into a low ability reading group, there is an increase in the achievement gap. Lleras and Rangel evaluated standardized reading scores and coded students based on the level of their respective reading group. Comparatively, minority students who were placed in a low level reading class performed significantly lower than students of the same academic ability who were placed in a classroom of mixed abilities.

Hong, Corter, Hong, and Pelletier (2012) studied the impact that homogeneous grouping in reading has on kindergarteners. Over 20,000 kindergarteners were tested and observed in order to determine how much time was spend in homogenous groups and which group they participated in: high, middle, or low. Contrary to many other results, Hong et al. discovered that homogeneous grouping does in fact optimize learning for middle and low-performing students when ample time is spent with differentiated instruction.

In addition to Hong et al.'s study, Robinson (2008) discovered that all kindergarteners benefit from ability groups within the classroom. He differentiated between students in this study based on their ethnicity. While Caucasian kindergarteners saw .14 points of growth in reading per month when participating in small groups with children of similar ability, Hispanic students who primarily spoke Spanish at home displayed .14 more points of growth per month than the Caucasian students. Robinson concluded that Hispanic students being in a low level reading class may have an impact on their self-concept; however, when placed in a group with other Spanish speakers the deficiency became one of language rather than competency. Hong et al. (2012) as well as Robinson found that grouping students into small reading groups by ability had a positive academic impact on students.

Puzio and Colby (2010) conducted a meta analysis of studies regarding within-class ability grouping as it pertains to reading instruction They analyzed the studies and discovered that within-class ability grouping does increase academic growth in reading and therefore should be continued. They stated that further research needed to be conducted in order to determine the specific type of grouping that is most beneficial and the instructional practices that are most effective in small groups.

Academic Impact in Math

Several researchers found that ability grouping in math leads to higher academic achievement (Pierce et al., 2011; Soloman, 2007). Pierce et al. discovered that students participating in a gifted and talented program performed higher on a pre- and posttest in math when clustering was implemented. Clustering is the practice of placing students of a particular ability, typically 8-10, in a heterogeneous classroom but providing differentiation. Over this 2-year study Pierce et al. found that when clustering for math instruction was implemented, with differentiated lessons and following the gifted and talented program, not only did the gifted students excel, but so did the others in the classroom. Pierce et al. concluded that the gifted and talented students were given differentiated instruction with students who performed at a similar ability in conjunction with interacting with students of all abilities. Further, the practice of clustering the gifted students forced teachers to be more attentive to differentiation and providing instruction tailored to the students individual needs in math. Overall, Pierce et al. found that the practice of clustering in third grade math is beneficial not only to students who participate in the gifted and talented program but also for the rest of the students in the classroom.

Soloman (2007) found similar results as Pierce et al. (2011). Soloman looked qualitatively at ability grouping in math. She interviewed 18 boys and girls in the middle grades,

some who were in a top group and others who were in a low group. Looking thematically at the data, she discovered that the top set of students expressed that they were challenged, given high expectations, and enjoyed exploring the creativity of math. However, the low set of math students stated that they spent much of their time on facts and finding the correct answer. They further expressed that when working in groups they enjoy investigation during group work; however, they had difficulty expressing different ways to complete a problem. Soloman found that overall students in top math groups expressed creativity, high expectations, and participation, whereas students in low math groups discussed fact finding, rote work, and had a desire to just get the work done. Both Soloman and Pierce et al. found that leveled instruction in math may be beneficial for high achieving grouped students; however, there are negative responses from students in low math groups.

Impact of Ability Grouping on Student Self-Concept and Social Inequality

"Uppertracks? Man, when do you think I see those kids? I never see them. Why should I? Some of them don't even go to class in the same building with me. If I ever walked in to one of their rooms they'd throw me out before the teacher even came in. They'd say I'd only be holding them back from their learning." (Cottle, 1974, p. 24) Ollie Taylor clearly was impacted emotionally when placed into a lower-track at his school. With research stating that there is a positive correlation between self-esteem and academic achievement (Kususanto, Ismail, & Jamil, 2010), it appeared that Ollie had an uphill battle ahead of him.

Gallagher (2011) found that many teachers do not have the knowledge regarding the social and emotional development of students, especially of gifted and talented students. When discussing strategies to meet the needs of students, many teachers stated that the best form of

acceleration is within-class, subject specific intervention. Teachers expressed concerns regarding elitism when pulling gifted and talents students into their own classroom as well as the need for real-world settings and socialization. Gallagher concluded that while many teachers feel that ability grouping is a positive instructional strategy, there is a need for mixed-ability classrooms to encourage social development and provide an equitable environment.

Self-Concept

One of the primary concerns regarding homogeneous grouping is that it impacts the self-concept of students. Student self-concept and perception of education is directly linked to their academic success. Ability grouping, while providing tailored academic instruction, has also led to research showing that it has an impact on student self-concept. Research has shown that with groups performing at a higher level of achievement, self-concept scores are higher, particularly in mathematics (Reuman, 1986). However, one of the many criticisms of ability grouping is the negative impact that is often associated with self-concept. In 2001 British researchers Ireson, Hallam, and Plewis conducted a study regarding the self-concept of 13-14 year olds as it pertains to their specific ability group. This research studied over 3,000 students and disaggregated the data based on type of ability grouping (no ability grouping, limited, or "streaming"), gender, specific level, and previous self-concept scores. Overall, the data showed that students exposed to limited ability grouping have higher self-concepts (Ireson et al., 2001).

Conversely, Hong et al. (2012) assessed over 20,000 kindergarten pre and posttests as well as their social-emotional scores. They found that contrary to many other studies, there was not a negative impact on low performing students with regards to their social-emotional scores when ample reading instruction took place in a homogeneous group. He added that there was no negative effect on students performing on a middle or high level with regards to their social-

emotional scores in kindergarten. However, Hong et al. indicated that in order for there to be no negative impact on the social-emotional score it is imperative that there be ample time spent in the homogeneous group with differentiated instruction.

Chmielewski (2013) categorized students into three groups based on the type of math instruction they received: course-by-course tracking, within school tracking, and between school tracking. Course-by-course tracking is common in the United States and is synonymous with homogenous grouping. Within school tracking is tailored toward secondary students in which they are grouped for all subjects based on a particular track. In her study evaluating the Programme for International Students (PISA) 2003 data set in conjunction with self-concept scores regarding math she found that students who participated in within-school or between-school tracking had similar scores to many other studies: high-achievers have low self-concept scores and low achievers have high self-concept scores. Conversely, Chmielewski discovered that students who participated in course-by-course tracking had the opposite results: high achievers had high self-concept scores and low-achievers had low self-concept scores. Therefore, Chmielewski reported that students participating in homogeneous grouping or course-by-course tracking in math tend to have high self-concept scores in math if they are a high-achiever, but a low self-concept score if they are a low achiever. She concluded that the there were differences in self-concept scores regarding math between the different types of ability grouping.

Further evidence has been found in research conducted by Wouters et al. (2012) in a study of high school students that tracked their self-concept scores while being ability grouped. When students were dropped to a lower level, their initial self-concept increased, as they were now the "big fish." However, their academic achievement decreased. The research found that there was a positive correlation between academic self-concept and academic achievement.

Kususanto et al. (2010) conducted a study regarding self-esteem and the impact of ability grouping. Three hundred two students from four public secondary schools in Malaysia completed a self-esteem survey. The students were ability grouped into leveled classrooms. Overall, the high achieving students scored significantly higher than the low achieving students; however, there was also a significant difference regarding teacher perception and self-esteem. Students in low-level classes indicated that their teachers were controlling, whereas high achievers stated their teachers were supportive. Kususanto et al. concluded that there is a positive correlation between the teacher's perception of students and their self-esteem; therefore, teachers may have different effects on students based on their level, which may have impacted their self-esteem.

When completing a study in Germany, researchers found that there was no statistical difference between the self-concept scores of high school ninth graders and their grades (Trautwein et al., 2006). Further, Ireson et al. (2001) found that there is no overall link between students' view of academics and their type of ability grouping. Ireson et al. stated that, "When other variables are statistically controlled, there is no significant effect of the extent of ability grouping in the school as a whole" (p. 297). In addition to the study that was conducted by Ireson et al., Ireson and Hallam conducted yet another study in 2005 and found that in regards to mathematics, ninth graders provided data that indicated no significant relationship between the group they were placed in for instruction and their self-concept regarding this subject area. Overall, the research has shown that in some circumstances ability homogeneous ability grouping has had a negative impact on self-concept, and in other cases is has not had a significant impact.

Social Inequality

Braddock and Slavin stated, "Ability grouping is ineffective. It is harmful to many students. It inhibits development of interracial respect, understanding and friendship. It undermines democratic values contributes to a stratified society. There are effective and practical alternatives" (1992, p. 14).

Boler (2008) defined relational equity as students interacting with each other with respect and responsibility. In his study Boler observed heterogeneous math classrooms in which there were many differences in ability, ethnicity, and culture. Through 600 hours of observation, Boler discovered that through using different teaching strategies, the students in these mixed ability groups not only achieved higher but also learned to respect other cultures, levels, genders, and learned good behavior. Boler concluded that through the use of heterogeneous classrooms, students gain relational equity through differentiated teaching strategies.

Sorting people into groups based on achievement is not a normal practice of society and researchers concurred that placing students into ability groups increases social inequality (Kintz, 2011). Fendler and Muzaffar (2008) wrote an article in which they discussed the history of the bell curve. Within the realm of education if a true bell curve is used to determine test results or student grouping then there is always an element of failure. Historically, the bell curve ensures that most of the students are "average" or "normal" while the rest of the students are either excelling or failing. Fendler and Muzaffar stated that the goal of any educational endeavor is to help students learn and succeed. If the bell curve is used, there is an aspect of failure. Therefore, when students are grouped by their academic ability there is one of two constructs: social reconstruction or social competition. Those in favor of sorting or ability grouping believe that competition is good in schools as educators are preparing students for the real world. Others

believe that schools should reconstruct the real world and students should be working together with others from all different backgrounds, socioeconomic statuses, and abilities. Overall, Fendler and Muzaffar stated that when educators implement sorting they believe that the "average" student is "normal" and therefore create unnatural competition and failure is part of the initiative. In other words, if students are all high achievers in an ability grouped classroom and the bell curve is implemented some of the students will fail, whereas they would not fail in a heterogeneous class.

A stigma is attached to ability grouping and many believe that it is a social injustice to subject students to social inequality in schools. Catsambis (2011) conducted a study in which she followed students in kindergarten through grade 5. She found that there are gender discrepancies in both the high and low level reading groups. The information from the study included scores on a standardized assessment and student placement as well as student demographic data. The information collected was analyzed and gave evidence to the fact that in small group reading boys are underrepresented in the high level groups whereas girls are overrepresented. In low-level groups girls are underrepresented and boys are overrepresented. These statistics provided evidence to Catsambis (2011) that there are gender inequalities that occurred in placing over 20,000 kindergarten students into reading ability groups.

Researchers discovered that there is a significant impact on social comparison that students experience in regards to ability grouping (Ireson et al., 2001). The literature reviewed has shown that there are both negative impacts of ability grouping and no correlation at all between self-concept scores and students' view of their abilities and confidence in mathematics and science (Trautwein et al., 2006). However, this research as a whole has indicated that there is

statistical significance between a negative self-concept score and ability grouping in the subject area of reading (Chorzempa, 2006).

Kintz (2011) found that there was also a discrepancy between students in high ability groups and low ability groups. When reviewing the literature, Kintz discovered that overall students in high ability grouped classrooms make friends and tend to come from high socioeconomic backgrounds. Further, he noted that students from low ability grouped classrooms tend to have a negative stigma attached to them and when tracked at an early age, may be placed into a less rigorous educational path. Overall Kintz discussed the negative impact that ability grouping has on low-ability grouped students because they are typically stereotyped, receive less rigorous instruction, and become friends with those in their classroom who tend to have less motivation for success and growth.

Friendships have a large impact on academic achievement according to Flashman (2012). Flashman studied a school system with two diverse school settings; one school was small and rural and the other was large and urban. The two school settings were coded for the purpose of analysis; however, the process for evaluating the schools was the same. As a result, Flashman discovered that high-achieving students were friends with other high achievers in the large, urban schools. The same was true for low achieving students. However, in the small, rural schools achievement level was not a significant determining factor in forming friendships. Flashman said that this might be due to the lower numbers of students or that there were not as many leveled classrooms. Further, the evaluation of test scores indicated that there was a correlation between academic achievement and future success. Based on this evidence Flashman concluded that friendships are often developed between two students with similar academic achievement, which also has an impact on future success.

Christie (2008) recounted a story of her son and a young student from Vietnam. The boys became fast friends when they were placed in a heterogeneous classroom and both excelled in math. However, when the boys were separated into different leveled classrooms, primarily because of the limited English proficiency, the two boys grew apart. The Vietnamese student was placed in classes that he called "baby" classes and on the vocational path; however, the other student was placed on the university track. Flashman's (2012) research is evident in this story in that friendships are formed based on academic achievement and have an impact on future success.

Ability grouping is a very common practice in the Asian community. Kemp and Watkins (2001) conducted a study focused on 10-11 year old students in Hong Kong. The sample consisted of 132 males and 148 females. Traditionally, students are broken into five differing bands, ability groups, at the end of primary school and sent to their respective "banded" school at the beginning of secondary school. According to Kemp and Watkins the band that is associated with each school is widely known by the surrounding population and families have felt much stigma attached with each respective school.

This study also asked students to rate their self-concept on a scale of one to five on a 76-question survey. An ANOVA was run to determine the significance of these results. Overall, the research showed that students performing at in high-level bands had lower selfconcept because they were constantly comparing themselves to other high performers. Conversely, the bands of lower-performing students had a higher self-concept because they were successful in their band. While these students have received instruction at a high level, research has shown a negative impact on the two extreme groups (Kemp & Watkins, 2001).

Although ability grouping allows teachers to meet the needs of a large group of students in the whole-group instructional setting, research showed that there are many cons to this type of grouping. Ireson et al. (2001) conducted a study comparing self-concept scores of 13 and 14 year olds between ability grouped schools and schools not grouped by ability. They found that students who have experienced the instructional setting not grouped by ability overall have a higher self-concept with regards to school. This adds research regarding lower achieving students having lower self-concept with ability grouping. Fram et al. (2007) stated there needs to be a balance between meeting the academic individual needs of students and the need for social equity.

Neihart (2007) compiled research about the socio-affective impact that ability grouping has on gifted students. While many researchers found that ability grouping has mixed results regarding self-concept, Neihart discovered that students who are gifted typically benefit from homogeneous grouping; however, there are students who benefit more from the socialization and interaction with peers of different abilities. In her review of literature she suggested that students who are gifted as well as those who perform at a lower ability may benefit emotionally and socially, but it may hurt students who are prematurely accelerated to an inappropriate level. In conclusion, Neihart found that students who are gifted tend to excel in a homogenous group socially, emotionally, and academically; however, each student must be evaluated individually as some function better in a heterogeneous classroom.

Assessment and Student Learning

There are many components that must be taken into consideration when analyzing student learning and assessment. From classroom assessments to standardized testing students

undergo many forms of assessment on a regular basis. However, Pekrun, Goetz, and Titz (2002) discovered that academic emotions have a large impact on students' motivation, learning strategies, cognitive resources, and academic achievement. Further, in their study Pekrun et al. found that when surveyed, students reported having both positive and negative emotions toward school and their academic experience. The most commonly reported academic emotion was anxiety, which is linked with a negative impact on the educational experience and in turn would have a negative impact on motivation, learning strategies, cognitive resources, and academic achievement. Therefore, it is important to analyze student learning and discover the most appropriate method of assessment.

History of High-Stakes Testing

High-Stakes testing or standardized testing has been part of the education realm since the Progressive period (1890-1920). The United States became focused on industry and tracking students in order to meet the needs of the current economy. Originally the standardized intelligence tests were created by the teachers and intended to place students into tracks according to their career outlook (Emery, 2007). However, standardized intelligence testing has morphed into high-stakes testing, which according to Emery is driven by corporate-inspired educational reform.

Within the past 30 years there has been a drive to standardize education as well as assessment in order to maintain high academic standards for American children. During the 1980s President Ronald Regan implemented three waves of education reform. The first wave had a goal to increase graduation requirements, standardize curriculum, and increase testing and certification requirements. The second wave focused on the local level of education and sought to increase professionalism and empower teachers. The third and final wave focused on children

and their learning. During the 80s there was an increase in the amount of involvement by the federal government and a push to standardize both curriculum and assessment. Emery (2007) argued that in 1989 at a Business Roundtable top CEOs of the county agreed to focus their attention on education reform through high-stakes testing.

President Clinton initiated education reform through Goals 2000 in which he focused on state content standards, high-stakes testing, and accountability for student achievement. With President George W. Bush's No Child Left Behind reform, educators saw another push for increasing rigor of standards, student assessment, teacher accountability, and teacher quality. However, through this time period there was a widening of the achievement gap; the high achieving students were making greater strides in achievement and the struggling students were being less successful. Therefore, President Obama decided to turn the focus from student achievement to student growth. Many states have passed legislation that link teacher evaluation with student growth measures rather than student achievement levels. While student achievement scores are reported and used to classify students into ability groups in many schools, teachers are more concerned with student growth. The student growth score is tied to the teacher evaluation in many states.

Defining Student Achievement

Standardized intelligence testing has morphed from using the data to place students appropriately into career tracks to an everyday occurrence through all grade levels. Standardized tests measure what is often referred to as student achievement. According to Oakes (2008) student achievement or academic achievement scores were derived out of a need for policy makers to monitor learners and accountability. While high-stakes testing results are not always accepted as consistent and precise, Stackett, Borneman, and Connelly (2008) found that some

standardized assessments do portray accurate results of student learning, but they stated that most standardized assessments do not reflect what a student has learned.

Armstrong (2006) differentiates between Academic Achievement Discourse and Human Development Discourse. He ascertained that many schools are focused on Academic Achievement Discourse or raising test scores and making instructional decisions based on what will help increase scores. According to Armstrong, however, there are many negative impacts of following the Academic Achievement Discourse such as implementing instructional practices that are developmentally inappropriate for students at a particular age and causing unnecessary and unhealthy stress for both teachers and students. Armstrong advocates following the Human Development Discourse, or teaching the whole child. He quoted Piaget in regards to the purpose of education stating, "To train young people to think for themselves and not to accept the first idea that comes to them." (Armstrong, 2006, p. 7) Therefore Armstrong advocates that teachers teach the whole child rather than conforming to a strict and uniform curriculum driven by achievement scores and focus on appropriately developing the entire child.

Measuring Student Learning

"Educational and psychological testing and assessment are among the most important contributions of behavioral science to our society . . . . There is extensive evidence documenting the effectiveness of well-constructed tests for uses supported by validity evidence" (AERA, APA, & NCME, 1999, p. 1). According to Stackett et al. (2008) there are many standardized tests that appropriately measure student learning. However, the Finnish education community would say that measuring student learning does not have to be assessed through high-stakes testing (Sahlberg, 2007).

Finnish students do not take a standardized test until they finish upper-secondary school. They believe that, "Primary school is, to a large extent, a 'testing-free zone' reserved for learning to know, to do, and to sustain natural curiosity. Teachers also experience more genuine freedom in curriculum planning; they do not need to focus on annual tests or exams" (Sahlberg, 2007, p.156) Through this freedom, teachers are able to focus their attention on providing students with basic knowledge in their first 6 years of education and move into reasoning and problem solving in the following years. The Finnish have adopted the following principals that outline their beliefs regarding education:

1. Some basic school for all students
2. Well-trained teachers – Almost all Finnish teachers have a master's degree
3. Intelligent accountability
4. Culture of trust
5. Sustainable leadership

While the Finnish education community has not latched on to the world-wide push for high-stakes testing, they have increased their student academic accountability through international assessments such as The International Mathematics and Science Study (TIMSS) and the Programme for International Assessment (PISA). Finland has not implemented high-stakes testing within its academic curriculum and has seen an increase in student success based on these two assessments as well as an increase in students completing their tertiary education (Sahlberg, 2007).

Stackett et al. (2008) and Popham (2011) found differing results as to the validity and accuracy of standardized tests as they pertained to student learning. Popham discovered that there are very few, if any, standardized tests that do an adequate job translating student learning

to a test score. Popham ascertained that while these high-stakes tests do not produce accurate reflections of student learning, teachers must be assessment literate. There is current research that discovered the use of formative assessment to modify instruction might produce up to double the student learning. He recommended that standardized testing, while used to monitor school growth and success as well as accountability for funding, should be replaced by an evaluation method that produces a more accurate reflection of student learning (Popham, 2011).

There are many concerns regarding the results of high-stakes testing and how accurately it measures student learning. Stackett et al. (2008) found that there are eight common beliefs regarding testing. While these beliefs are widely accepted, there is contradicting nation-wide research regarding each of the statements listed below.

1. Tests are poor predictors
2. Tests measure short-term learning
3. They are used for predictive value
4. Tests reflect socioeconomic status rather than developed abilities
5. Tests are coached
6. Tests are biased against minority groups
7. Minority groups tend to score lower on tests but perform just as well as the majority group when hired for a job
8. Motivation or threat are the reason for minority group mean differences

After pooling and analyzing large amounts of data from previous studies, Stackett et al. discovered that if tests are constructed correctly, they can accurately portray student learning. The following results were found from their study (Stackett et al., 2008):

1. High-stakes tests are generally valid regarding the correlation between student ability and test results.
2. Validity is not a result of socioeconomic status.
3. Coaching is not a factor of performance on testing.
4. High-stakes tests do not tend to underscore minority students.
5. Motivating students prior to testing does not have an impact on performance.

Stackett et al. discovered that there is a positive correlation between student ability and standardized test scores; however, they did caution educators to make sure that the test is valid and reliable, avoiding bias at all costs. Not all standardized tests are capable of providing an accurate measure of student learning. They stated that some tests that claim to be valid can be used inappropriately, causing them to extinguish their validity.

Limits of High-Stakes Testing

The government implemented high-stakes testing as a way to monitor and track student growth and the quality of teacher instruction (Popham, 2007). Lee (2008) pointed out that based on a meta-analysis of 14 studies, high-stakes testing policy had zero effects on the achievement gap. He stated that there is still a large racial achievement gap that is growing rather than shrinking. Lee (2008) discovered that the evidence that high-stakes testing is reliable was limited and that further study needed to be conducted in order to determine the extent to which high-stakes testing policies impact student achievement and the closure of the achievement gap.

Popham (2007) discussed the implications of an instructionally sensitive versus an insensitive standardized assessment. An instructionally sensitive test is given and scores are associated with quality teachers; however, the scores are not a direct result of quality teaching alone. Instructionally insensitive tests do not allow for interpretation of scores as it pertains to

either strong or weak instruction. Popham (2007) provided several features of instructionally sensitive tests:

1. Number of Instructional Aims: The more instructional aims present on an assessment, the less sensitive it becomes.

2. Clarity of Assessment Targets: Teachers must understand what they are to teach their students.

3. Items Assessed Per Curricular Aim: Teachers need to understand which aims correlate with assessment questions so that they can evaluate their own instruction and use it as a tool for improvement.

4. Item Sensitivity: An item is sensitive when the following statements are true: answering the item is not influenced by SES or inherited aptitude but rather is a response to instruction.

When an assessment endures this evaluation and is deemed sensitive, Popham (2007) stated that it is a sensitive test and therefore can be used to measure quality instruction. However, Popham found that current high-stakes tests do not meet the criteria set out, and therefore are insensitive and a poor measure of teacher instruction and student learning.

Further research conducted by Riffert (2005) explored the topic of standardized testing. Riffert claimed that due to the accountability teachers were forced to teach to the test and have lost their creativity and flexibility to teach real world, hands-on, creative tasks. Zirkel (1998) stated that a Kentucky student was denied graduation and his sister was denied promotion to the next grade because their family refused to allow them to take the statewide standardized assessment. The students were not denied their promotions due to learning but rather to state-

mandated testing. When they did not take the text, they received a “0”, causing them to not graduate or be promoted (Zirkel, 1998).

Guisbond and Neill (2004) studied the implications of No Child Left Behind legislation. They discovered that there were two main suppositions regarding No Child Left Behind. First, boosting test scores should be the primary goal of schools. This assumption indicates that education can have a one-size-fits-all approach. It also does not take into consideration the social goals of schools and differentiating instruction in order to meet the needs of the students (Guisbond & Neill, 2004). Second, as a result of poor teaching student performance is low and can be corrected by threats and sanctions. This does not touch the root of the problem. In order to increase student achievement, teacher quality must be increased, rather than limiting teachers to teach to the test. Although schools do not receive Adequate Yearly Progress (AYP) anymore, the following statement holds true regarding Annual Measureable Objectives (AMOs): "The AYP requirement, a completely arbitrary mathematical function grounded in no defensible knowledge or theory of school improvement, could, and probably will, result in penalizing and closing schools that are actually experts in school improvement" (Elmore 2003, p. 7).

Guisbond and Neill (2004) continued, stating that elements of a better accountability system is:

1. Get federal, state, and local governments to work together to provide an opportunity for all students to learn a rich curriculum.
2. Use multiple forms of evidence to assess student learning rather than a single standardized assessment.
3. Help teachers and schools to promote success for all students.

4. Accountability should be more localized, including local school boards, educators, parents, and members of the local community.
5. Focusing the state government on providing support and tools for schools to promote equity in schools and teaching.

Guisbond and Neill ascertained that if these principles were implemented into the legislative practices regarding education students would have the opportunity to learn and not be limited based on the scores they earned on a standardized assessment. Student learning should be assessed in progress evaluations and the amount of standardized testing should be reduced according to Guisbond and Neill (2004).

Summary

Much of the literature surrounding ability grouping is conflicting. Betts (2011) completed a comprehensive literature review of these studies both within the United States and abroad. Betts found that while there is conflicting evidence that both support and refute ability grouping, the discrepancy is not necessarily in the methodology but rather the differences in definitions of ability grouping and poor measurements. There are many terms associated with ability grouping: banding, tracking, streaming, setting, homogeneous, heterogeneous, ability grouped, and not grouped by ability. However, Betts concluded that with regards to self-concept and social inequality, in the United States grouping students by ability or not grouping by ability does not impact either of these two factors. Internationally, there is evidence that proves that that inequality does exist within the realm of ability grouping. Betts attributed this to the difference in resources available. Overall, Betts concluded that research regarding ability grouping is difficult to compare because the results are specific to the country, region, and demographics.

Good and Marshall (1984) found similar results to Betts (2011). They found that the research of ability grouping is so varied primarily due to the extent of homogeneity or heterogeneity within classrooms. While schools may in fact place students into classrooms based on academic level, there may be one school included in a study that has 15 gifted and talented students whereas another school has none. Good and Marshall ascertained that there are differences in quantitative results because of the inconsistency of classroom makeup.

The extant literature shows that there are both benefits and drawbacks to grouping students homogenously or heterogeneously by ability; however, it is clear that grouping students into small groups allows the teacher to meet specific needs of students and differentiate instruction. While heterogeneous grouping provides a real-life setting in which students work with others that have differing abilities, homogenous grouping allows the teacher to tailor instruction to the needs of the entire class and has shown to improve student achievement scores in some studies. While differentiation becomes more necessary in a heterogeneous classroom, students in homogeneous groups may struggle with self-concept. Cheng, Lam, and Chan (2008) emphasized that within small groups it is important to have quality group processes. Heterogeneity was not a determining factor in relation to student self-efficacy for both low and high groups, but rather the group processes benefited the overall collaborative learning experience. Therefore, it is important to ensure that regardless of the composition of the small group the processes and purpose be of quality.

CHAPTER 3

METHODOLOGY

Introduction

This chapter is focused on the methodology used in this nonexperimental, quantitative study to evaluate if there was a statistically significant difference in percentages of students scoring in each proficiency category on the TCAP test in reading and math between ability grouped classrooms and classrooms not grouped by ability. Research provided evidence that both supports and refutes the practice of ability grouping. This study is focused on academic achievement proficiency percentages on the Tennessee Comprehensive Assessment Program (TCAP) as reported on the Tennessee Department of Education website for schools in rural East Tennessee. The results were limited to grades 3, 4 and 5 in the subject areas of reading and math for the 2012- 2013 school year. Comparisons were made in the subjects of reading and math as well in grades 3, 4, and 5 with regards to grouping practices: ability grouped and classrooms not grouped by ability. Schools were coded based on information collected from school principals identifying grouping practices.

Research Questions and Corresponding Null Hypotheses

The following research questions were addressed in this study. The independent variables were the type of grouping: ability grouped or not grouped by ability. The dependent variables were the academic achievement proficiency percentages on the TCAP in reading and math.

Research Question 1

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic,

Below Basic) for third grade math between schools that have students grouped by ability and schools that have students not grouped by ability?

H_o1_1: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the below basic level for third grade math between schools that have students grouped by ability and schools that do not group students by ability?

H_o1_2: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the basic level for third grade math between schools that have students grouped by ability and schools that do not group students by ability?

H_o1_3: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the proficient level for third grade math between schools that have students grouped by ability and schools that do not group students by ability?

H_o1_4: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the advanced level for third grade math between schools that have students grouped by ability and schools that do not group students by ability?

Research Question 2

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for fourth grade math between schools that have students grouped by ability and schools that do not group students by ability?

H_o2_1: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the below basic level for fourth grade math between schools that have students grouped by ability and schools that do not group students by ability?

H_o2_2: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the basic level for fourth grade math between schools that have students grouped by ability and schools that do not group students by ability?

H_o2_3: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the proficient level for fourth grade math between schools that have students grouped by ability and schools that do not group students by ability?

H_o2_4: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the advanced level for fourth grade math between schools that have students grouped by ability and schools that do not group students by ability?

Research Question 3

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for fifth grade math between schools that have students grouped by ability and schools that do not group students by ability?

H_o3_1: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the below basic level for

third grade math between schools that have students grouped by ability and schools that do not group students by ability?

H_o3_2: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the basic level for fifth grade math between schools that have students grouped by ability and schools that do not group students by ability?

H_o3_3: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the proficient level for fifth grade math between schools that have students grouped by ability and schools that do not group students by ability?

H_o3_4: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the advanced level for fifth grade math between schools that have students grouped by ability and schools that do not group students by ability?

Research Question 4

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for third grade reading between schools that have students grouped by ability and schools that do not group students by ability?

Ho_o4_1: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the below basic level for third grade reading between schools that have students grouped by ability and schools that do not group students by ability?

H_o4_2: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the basic level for third grade reading between schools that have students grouped by ability and schools that do not group students by ability?

H_o4_3: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the proficient level for third grade reading between schools that have students grouped by ability and schools that do not group students by ability?

H_o4_4: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the advanced level for third grade reading between schools that have students grouped by ability and schools that do not group students by ability?

Research Question 5

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for fourth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

Ho_o5_1: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the below basic level for fourth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

Ho_o5_2: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the basic level for fourth

grade reading between schools that have students grouped by ability and schools that do not group students by ability?

H_o5_3: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the proficient level for fourth grade reading between schools that have students grouped by ability and schools that have students not grouped by ability?

H_o5_4: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the advanced level for fourth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

Research Question 6

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for fifth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

H_o6_1: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the below basic level for fifth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

H_o6_2: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the basic level for fifth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

H_o6_3: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the proficient level for fifth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

H_o6_4: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the advanced level for fifth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

Instrumentation

When collecting data for quantitative analysis it was imperative to choose an instrument that has established reliability and validity (MacMillian & Schumacher, 2011). This study incorporated both primary and secondary sources of data. School data were collected from the Tennessee Department of Education website, which provides yearly report cards. The achievement data collected were the proficiency percentages for each school for grades 3, 4, and 5 in both reading and math. The state reported the percentages in the following categories: advanced, proficient, basic, and below basic. McMillan and Schumacher emphasized that standardized tests provide reliability on a large-scale and have results that are objective. The Tennessee Comprehensive Assessment Program (TCAP) is described by the Tennessee Department of Education. "The TCAP Achievement test is a timed, multiple choice assessment that measures skills in reading, language arts, mathematics, science and social studies. Tennessee is an English-only state, which means all assessments are provided in English only. We *(the state)* do not offer any of our assessments in a native language format." (TDOE, 2014, paragraph 1).

Population and Sample

To control for socioeconomic or cultural differences, the population for this study was defined as rural public schools in East Tennessee. There are 33 counties that comprise East Tennessee, some which are not rural. These counties house approximately 350 public schools. Stratified sampling was implemented because not all schools fall into the category of either ability grouped or not grouped by ability classroom settings. The sample size was dependent upon the number of schools that implement either ability grouped or classrooms not grouped by ability for reading and math. Lack of response from directors of schools meant that not all schools were contacted. Therefore, this study collected data for approximately 20 schools that group students by ability and 30 schools that group students with varied ability in each classroom.

Data Collection

The data were collected from the Tennessee Department of Education School Report Cards website. To ensure reliability and validity, standardized test scores were used. McMillan and Schumacher (2010) indicated that standardized tests provide uniformity in administration and scoring; therefore, they provide student achievement scores that are consistently administered across counties. When collecting data regarding the type of ability grouping implemented, the researcher contacted a school administrator at each school to ensure that there is clear communication and understanding and accurate data were collected.

To code the schools as implementing either ability grouped or classrooms not grouped by ability, the researcher contacted individual principals. Principals were contacted through school email. The researcher defined ability grouped and classrooms not grouped by ability in the email and each administrator was asked to code each grade level (3, 4, and 5) and subject (reading and

math) as either ability grouped classrooms or classrooms not grouped by ability in the 2012-2013 school year. If principals did not respond within 1 week, the researcher followed-up with phone calls. For the school data to be used, the grade level as a whole had to apply either ability-grouped classrooms or classrooms not grouped by ability in the 2012-2013 school year.

The data collected for this study were loaded into IMB - SPSS data files and steps were followed as outlined by Green and Salkind (2011). When elementary schools were coded, the researcher accessed Tennessee Comprehensive Assessment Program data located on the Tennessee Department of Education Website.

Human Safety

To ensure human safety within this study the Institutional Review Board reviewed the application prior to research being conducted. The Institutional Review Board determined this not to be research with human subjects.

Validity, Reliability, and Bias

The researcher must take every measure to ensure that the study is valid and reliable, and avoid bias. Reliability and validity must be addressed during each step of the research process. When designing the research four types of design validity should be addressed; statistical conclusion validity, internal validity, construct validity, and external validity (McMillan & Schumacher, 2010). The data collection process should be reviewed in order to evaluate accuracy. When collecting information from school administrators, it was imperative to ensure that correct data were received in order to avoid incorrect coding. In analyzing the data committee members were consulted in order to ensure that the correct test was run and that the

data were read correctly. Further, the data were crosschecked in order to ensure that they have been imputed accurately, producing a coefficient of agreement (McMillan & Schumacher, 2010).

Researcher bias can be found at any point during the research process. In order to avoid bias, the researcher sought review from committee members. Furthermore, the researcher consulted experts in order to avoid sampling bias, which could gravely impact this study (McMillan & Schumacher, 2010).

Ethical Issues

In considering the ethical issues surrounding this study, it is clear that transparency is imperative (McMillan & Schumacher, 2010). All procedures and data are present in the written report so as to provide all information. Because the data collected for this study were public, privacy and confidentiality were not a concern. However, honesty and accurately representing the data were top priorities. Further, taking measures to avoid researcher bias was essential when conducting this study.

Data Analysis

Prior to running the data through IBM - SPSS the data were entered into a spreadsheet (McMillan & Schumacher, 2010). Independent samples t-tests were run in order to assess the relationship between school achievement scores (basic, below basic, proficient, and advanced) and the type of grouping. The alpha level was set at .05. Further, independent samples t-tests were run to determine if there is a statistical relationship in percentages between ability grouped and nonability grouping among grades 3, 4, and 5 in both reading and math.

Summary

The purpose of this study was to determine if there is a statistically significant difference between school-wide student achievement in grades 3, 4, and 5 based on the type of ability grouping: ability grouped or not grouped by ability classroom settings. Schoolwide TCAP data were collected form the Tennessee State Department of Education website for the 2012-2013 school year for grades 3,4, and 5 in reading and math. Public schools in East Tennessee were contacted by the researcher to determine if students were grouped by ability for reading and math or if they were placed in heterogeneous classrooms for instruction. Independent samples t-tests were conducted to determine the significance between academic achievement and the type of grouping in grades 3, 4, and 5 in reading and math.

CHAPTER 4

ANALYSIS OF DATA

The purpose of this study was to determine if there is a statistically significant difference between school-wide student achievement in grades 3, 4, and 5 based on the type of ability grouping: ability grouped or not grouped by ability. Percentages of students scoring in each of the proficiency categories (below basic, basic, proficient, and advanced) were divided by grade level and subject area: reading or math. To evaluate if there was a significant difference between schools that group students by ability and schools that do not groups students by ability, independent samples t-tests were run. The data were collected from public schools in rural East Tennessee. TCAP achievement scores were analyzed for grades 3, 4, and 5 in both reading and math. All TCAP data were accessed from the Tennessee Department of Education website and data analyzed were from the 2012-2013 school year.

In this chapter data were analyzed with regards for six research questions and 24 null hypotheses. Data were collected for schools that either grouped students for reading or math by ability or schools who did not group students by ability in the 2012-2013 school year. Two data measures were analyzed: percentages of students scoring in each proficiency level on the 2013 TCAP in math and reading and the type of grouping: ability grouped or not grouped by ability. The data were obtained from the Tennessee Department of Education website as well as from individual school administrators.

Analysis of Research Questions

Research Question 1

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic,

Below Basic) for third grade math between schools that have students grouped by ability and schools that do not group students by ability?

H_o1_1: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the below basic level for third grade math between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in math of third grade students scoring in the below basic category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(54) = 1.05$, $p = .297$. Therefore, H_o1_1 was retained. The percentage of students scoring in the below basic category in ability grouped classrooms ($M = 6.27$, $SD = 5.53$) was similar to those who were in a classroom not grouped by ability ($M = 7.97$, $SD = 5.41$). The 95% confidence interval for the difference in means was -4.93 to 1.53. The η^2 index was .020, which indicated a small effect size. Figure 1 shows the distributions for the two groups.

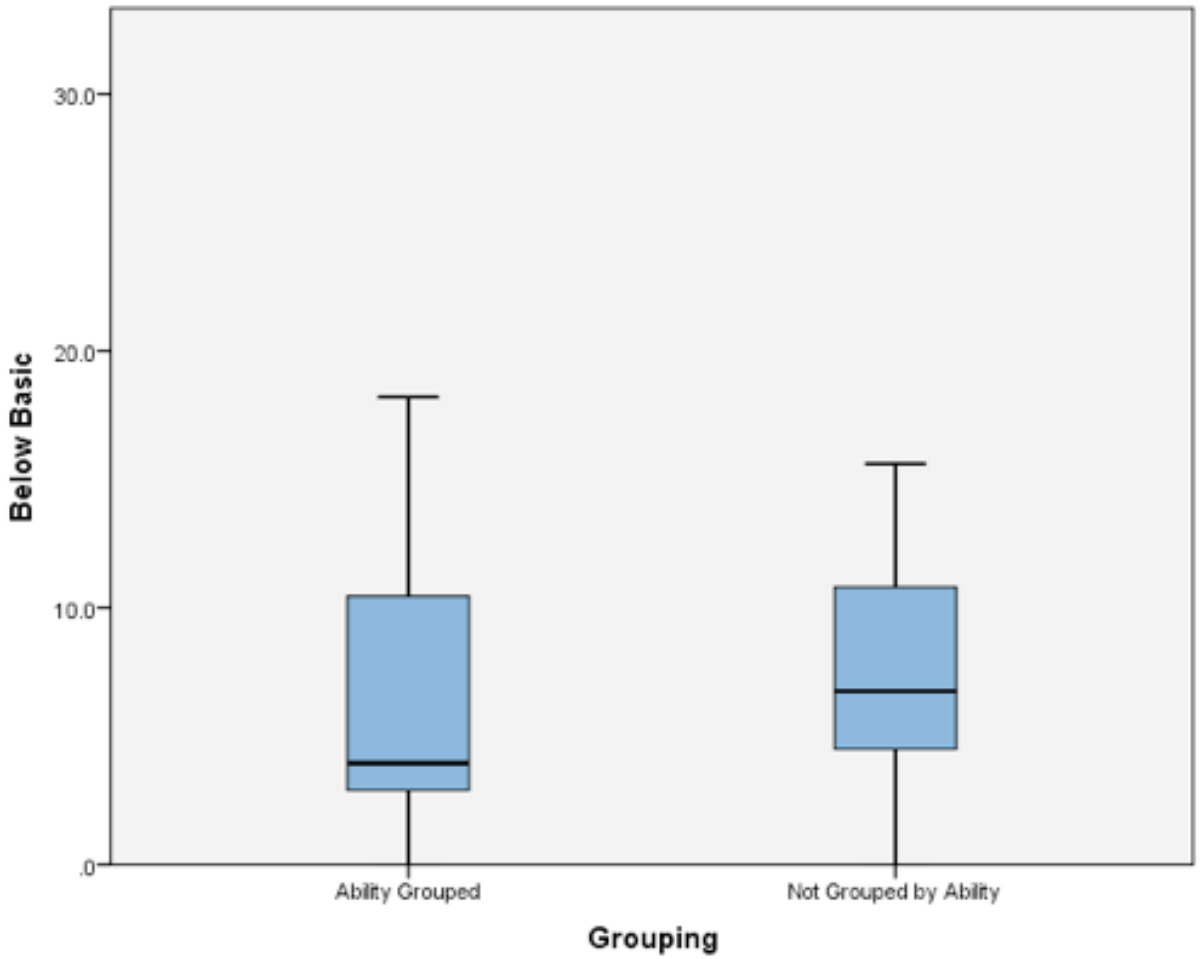

Figure 1: 3rd Grade Math: Percentage of Students Scoring in the Below Basic Proficiency Level

H_o1_2: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the basic level for third grade math between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in math of third grade students scoring in the basic category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped

by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(54) = .83$, $p = .411$. Therefore, H_o1_2 was retained. The percentage of students scoring in the basic category in ability grouped classrooms ($M = 33.68$, $SD = 13.59$) was similar to those who were in a classroom not grouped by ability ($M = 36.55$, $SD = 10.91$). The 95% confidence interval for the difference in means was -9.81 to 4.08. The η^2 index was .013, which indicated a small effect size. Figure 2 shows the distributions for the two groups.

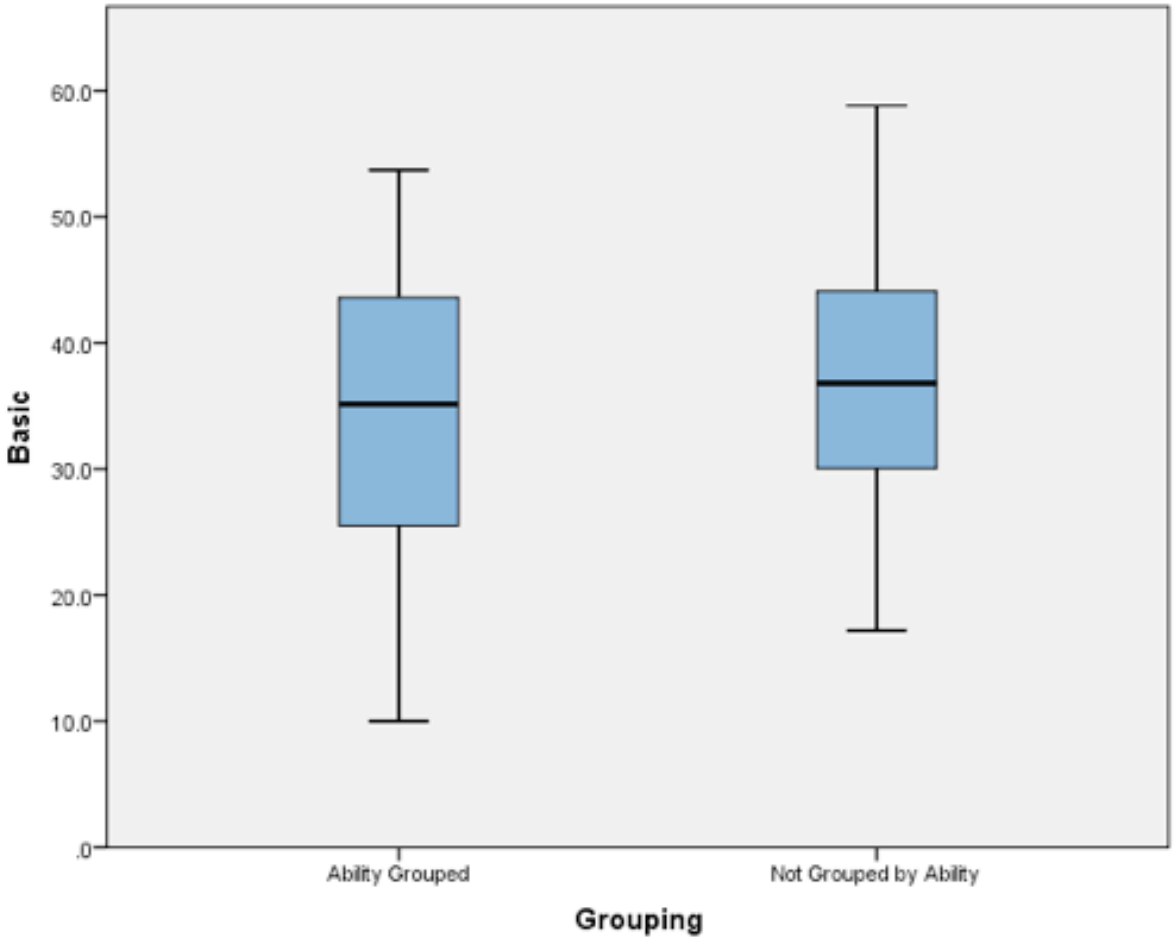

Figure 2: 3rd Grade Math: Percentage of Students Scoring in the Basic Proficiency Level

H_o1_3: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the proficient level for third grade math between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in math of third grade students scoring in the proficient category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(54) = 1.67$, $p = .100$. Therefore, H_o1_3 was retained. The percentage of students scoring in the proficient category in ability grouped classrooms ($M = 41.21$, $SD = 10.66$) was similar to those who were in a classroom not grouped by ability ($M = 36.79$, $SD = 8.19$). The 95% confidence interval for the difference in means was -0.87 to 9.93. The η^2 index was .049, which indicated a medium effect size. Figure 3 shows the distributions for the two groups.

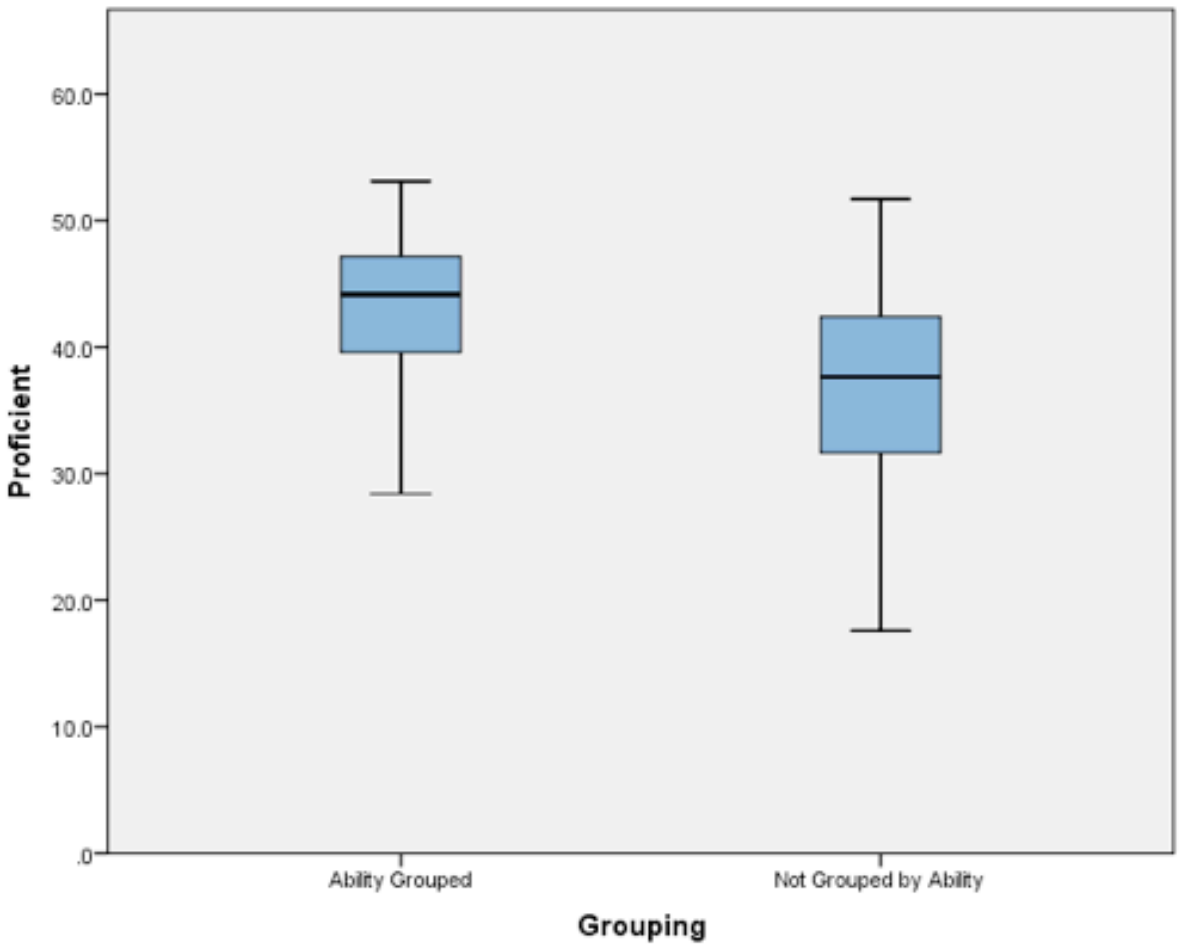

Figure 3: 3rd Grade Math: Percentage of Students Scoring in the Proficient Proficiency Level

H_o1_4: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the advanced level for third grade math between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in math of third grade students scoring in the advanced category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped

by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(54) = .05$, $p = .963$. Therefore, H_o1_4 was retained. The percentage of students scoring in the advanced category in ability grouped classrooms ($M = 18.34$, $SD = 9.92$) was similar to those who were in a classroom not grouped by ability ($M = 18.70$, $SD = 10.09$). The 95% confidence interval for the difference in means was -5.82 to 6.10. The η^2 index was $<$ .001, which indicated a small effect size. Figure 4 shows the distributions for the two groups.

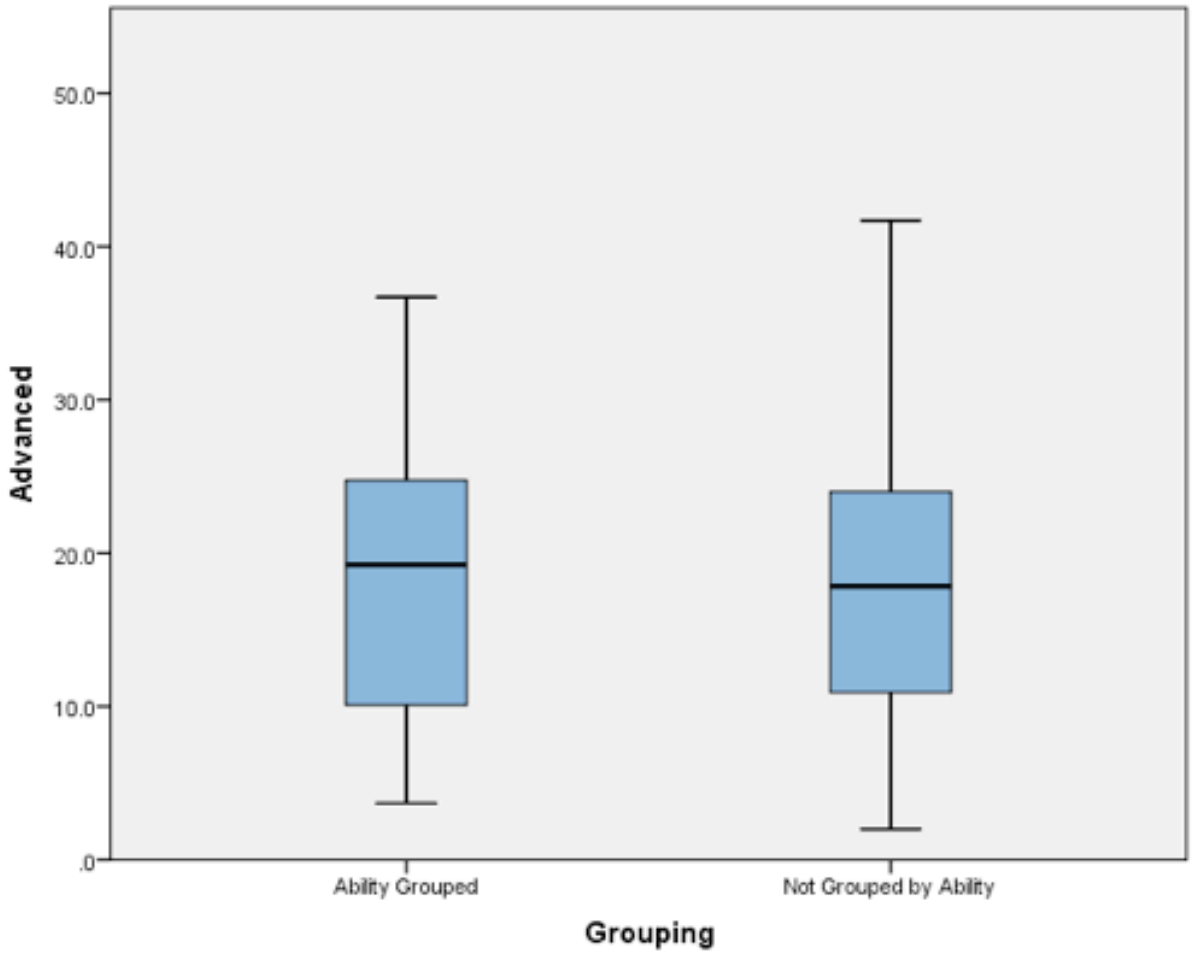

Figure 4: 3rd Grade Math: Percentage of Students Scoring in the Advanced Proficiency Level

Research Question 2

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for fourth grade math between schools that do not group students by ability?

H_o2_1: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the below basic level for fourth grade math between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in math of fourth grade students scoring in the below basic category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(53) = .17$, $p = .867$. Therefore, H_o2_1 was retained. The percentage of students scoring in the below basic category in ability grouped classrooms ($M = 14.35$, $SD = 14.38$) was similar to those who were in a classroom not grouped by ability ($M = 13.77$, $SD = 11.10$). The 95% confidence interval for the difference in means was -6.34 to 7.50. The η^2 index was $< .001$, which indicated a small effect size. Figure 5 shows the distributions for the two groups.

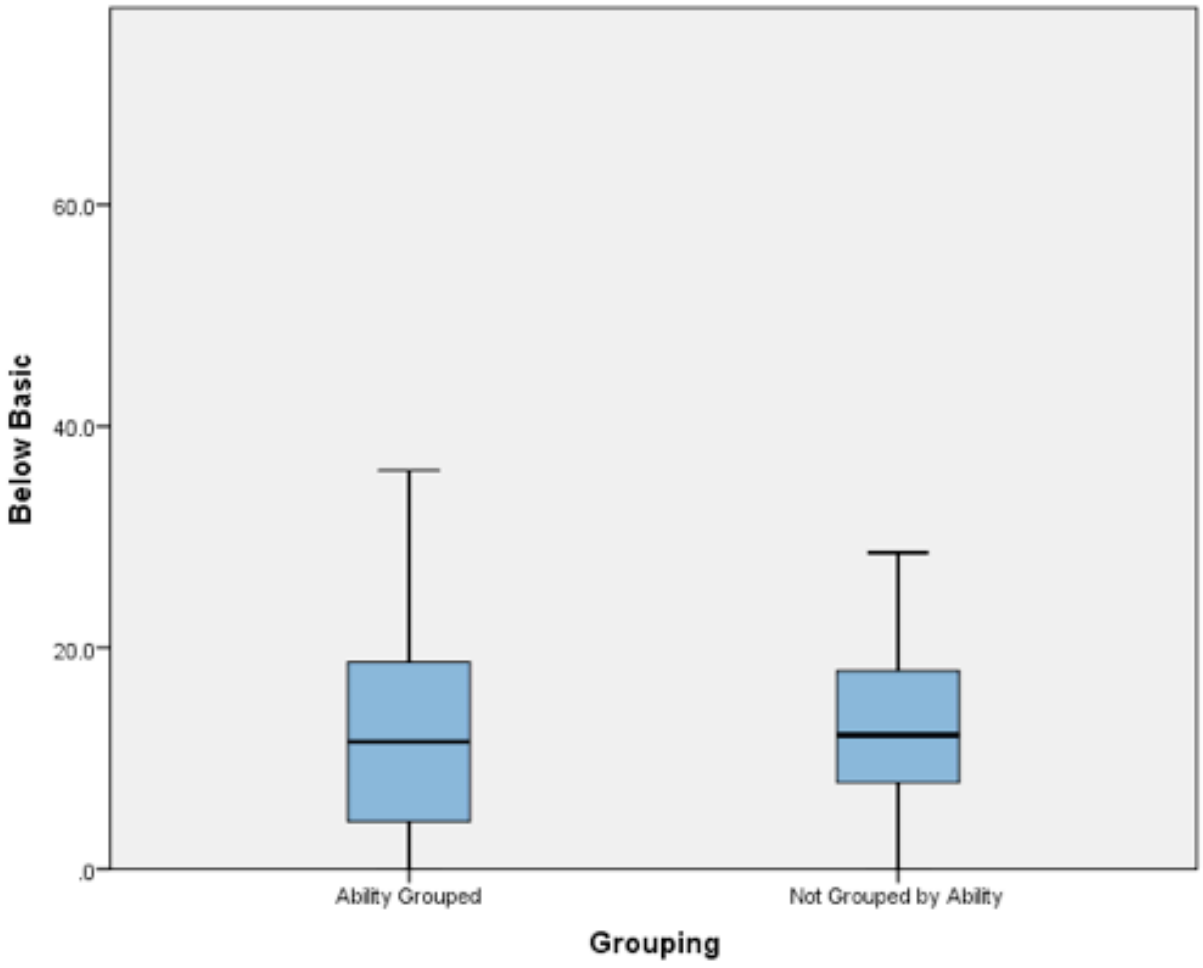

Figure 5: 4th Grade Math: Percentage of Students Scoring in the Below Basic Proficiency Level

H_o2_2: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the basic level for fourth grade math between schools that have students grouped by ability and schools that do not group students by ability.

An independent samples t-test was conducted to evaluate whether the mean percentages in math of fourth grade students scoring in the basic category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped

by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(53) = .55$, $p = .583$. Therefore, H_o2_2 was retained. The percentage of students scoring in the basic category in ability grouped classrooms ($M = 38.86$, $SD = 12.65$) was similar to those who were in a classroom not grouped by ability ($M = 37.12$, $SD = 12.65$). The 95% confidence interval for the difference in means was -4.59 to 8.07. The η^2 index was .006, which indicated a small effect size. Figure 6 shows the distributions for the two groups.

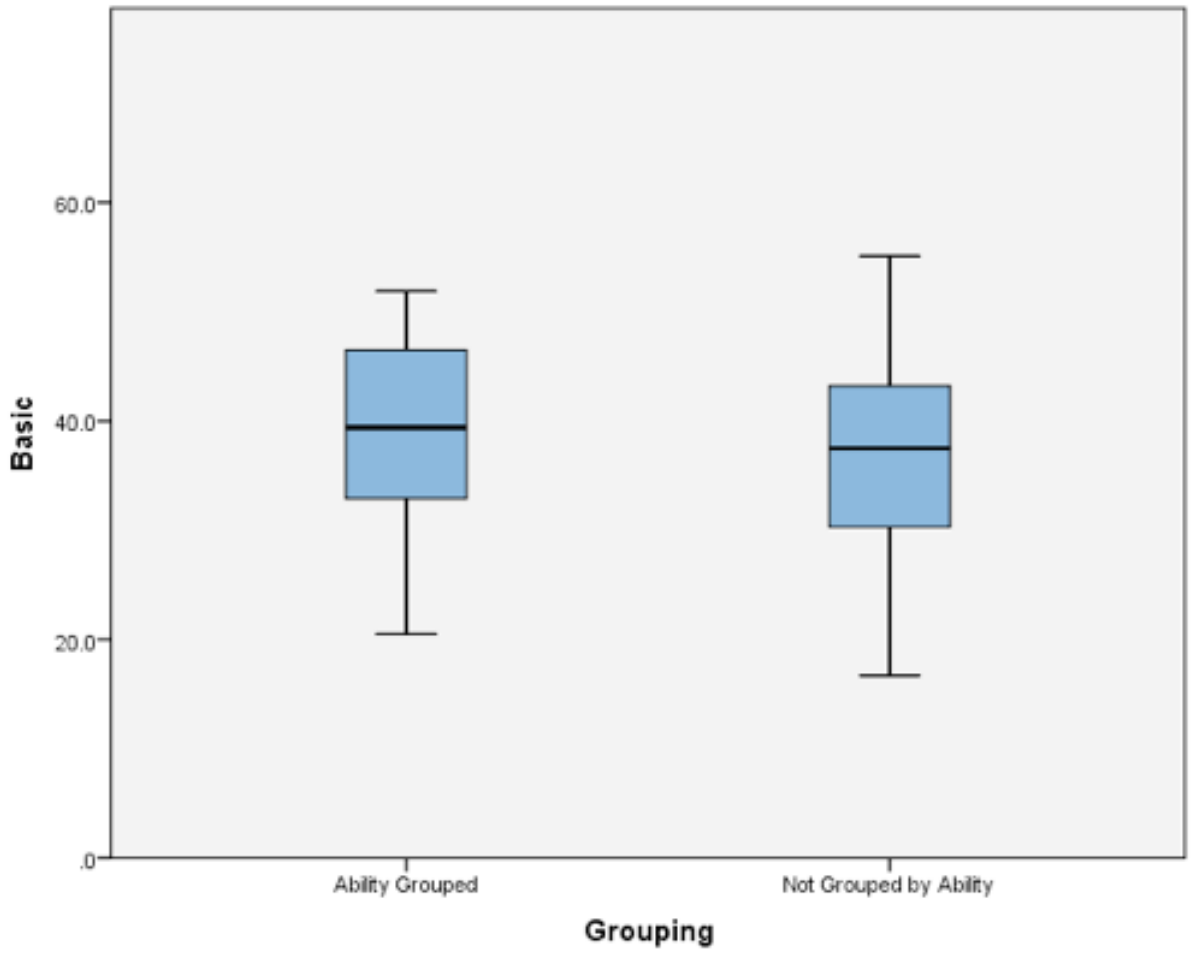

Figure 6: 4th Grade Math: Percentage of Students Scoring in the Basic Proficiency Level

H_o2_3: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the proficient level for fourth grade math between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in math of fourth grade students scoring in the proficient category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(53) = .13$, $p = .897$. Therefore, H_o2_3 was retained. The percentage of students scoring in the proficient category in ability grouped classrooms ($M = 32.93$, $SD = 14.44$) was similar to those who were in a classroom not grouped by ability ($M = 33.37$, $SD = 10.25$). The 95% confidence interval for the difference in means was -7.11 to 6.24. The η^2 index was < .001, which indicated a small effect size. Figure 7 shows the distributions for the two groups.

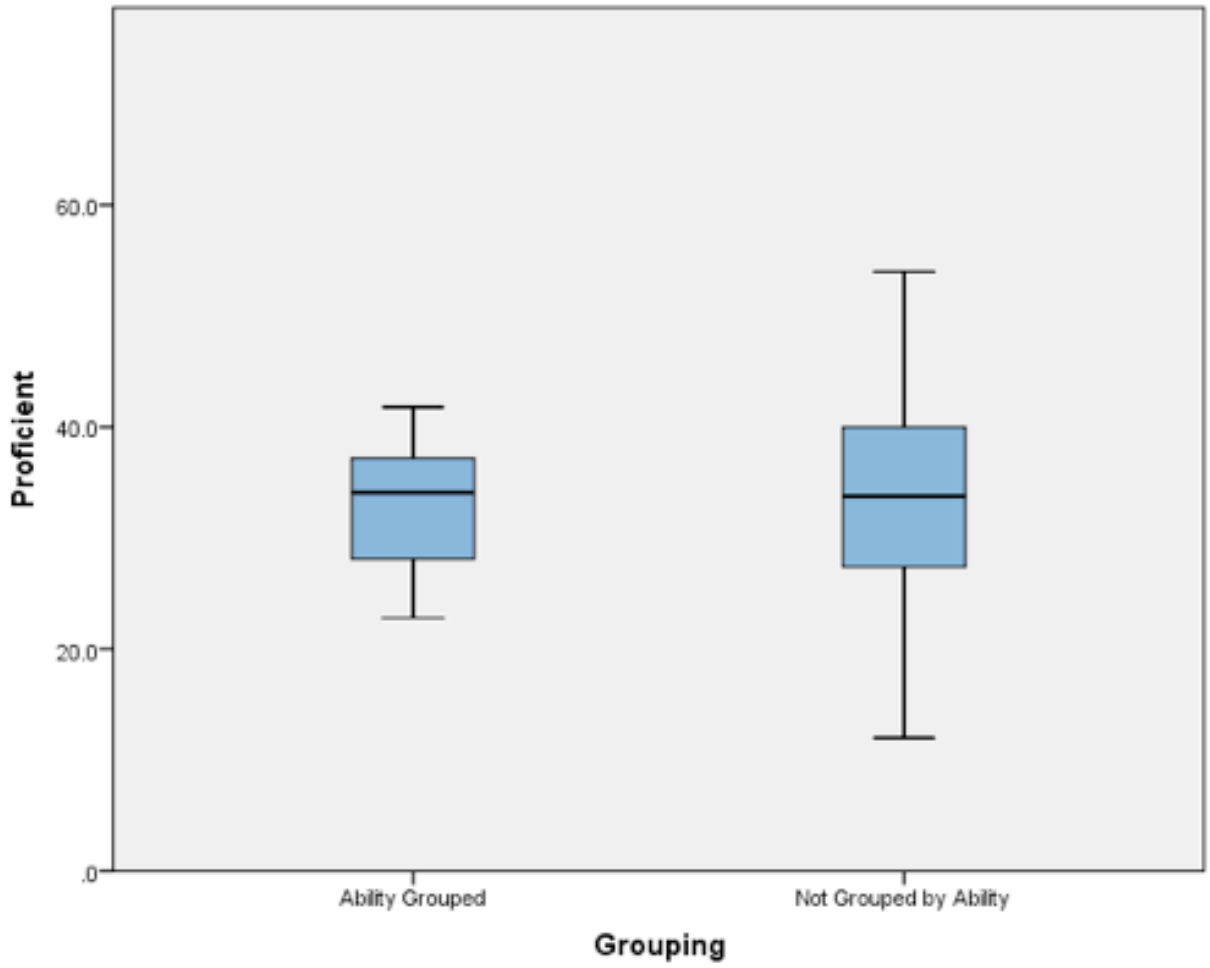

Figure 7: 4th Grade Math: Percentage of Students Scoring in the Proficient Proficiency Level

H_o2_4: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the advanced level for fourth grade math between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in math of fourth grade students scoring in the advanced category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not

grouped by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(53) = .79$, $p = .434$. Therefore, $Ho2_4$ was retained. The percentage of students scoring in the advanced category in ability grouped classrooms ($M = 13.60$, $SD = 7.90$) was similar to those who were in a classroom not grouped by ability ($M = 15.74$, $SD = 10.80$). The 95% confidence interval for the difference in means was -7.60 to 3.31. The η^2 index was .011, which indicated a small effect size. Figure 4 shows the distributions for the two groups.

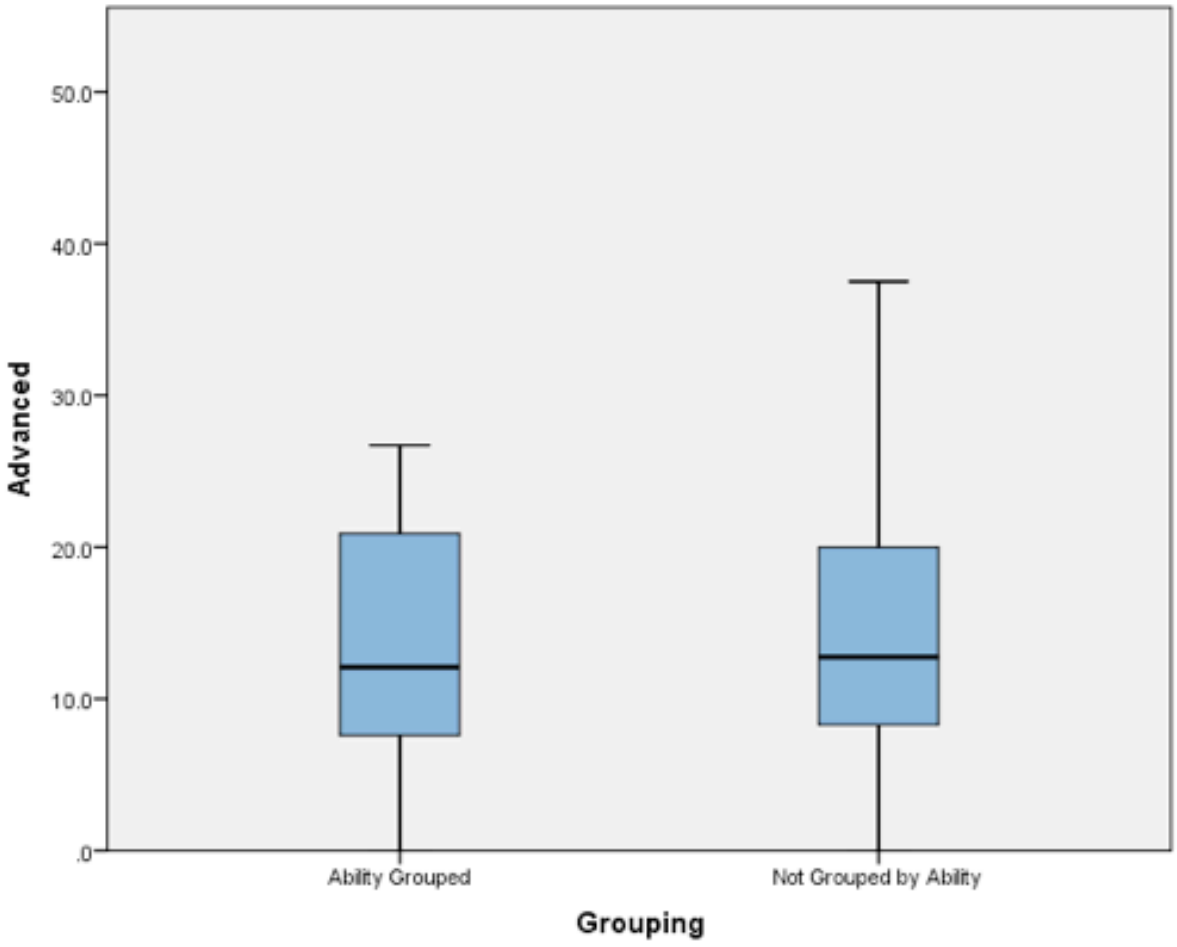

Figure 8: 4th Grade Math: Percentage of Students Scoring in the Basic Advanced Level

Research Question 3

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for fifth grade math between schools that have students grouped by ability and schools that do not group students by ability

$Ho3_1$: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the below basic level for fifth grade math between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in math of fifth grade students scoring in the below basic category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(50) = .55$, $p = .584$. Therefore, $Ho3_1$ was retained. The percentage of students scoring in the below basic category in ability grouped classrooms ($M = 15.90$, $SD = 15.12$) was similar to those who were in a classroom not grouped by ability ($M = 13.88$, $SD = 11.31$). The 95% confidence interval for the difference in means was -5.36 to 9.40. The η^2 index was .006, which indicated a small effect size. Figure 9 shows the distributions for the two groups.

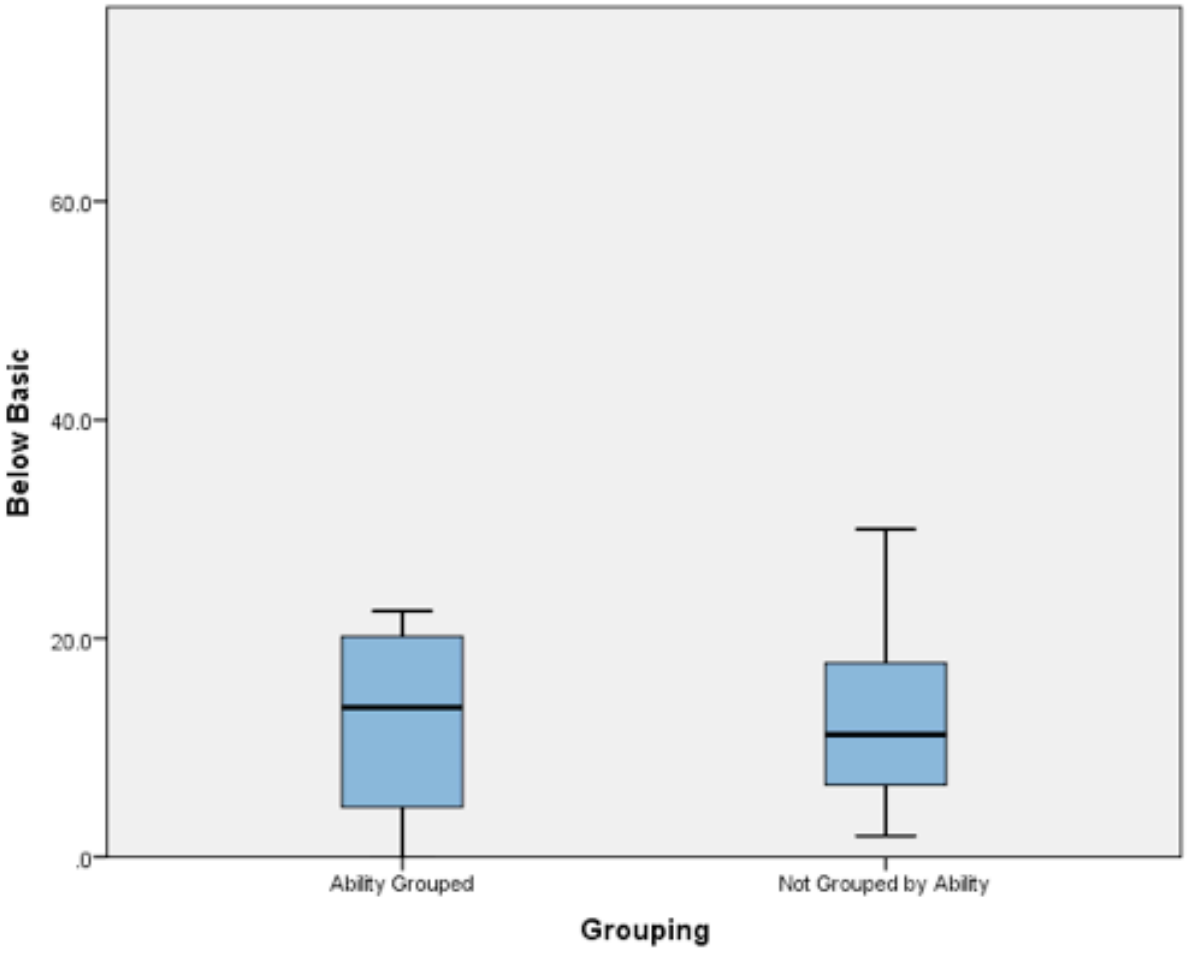

Figure 9: 5th Grade Math: Percentage of Students Scoring in the Below Basic Proficiency Level

H_o3_2: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the basic level for fifth grade math between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in math of fifth grade students scoring in the basic category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped

by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(50) = .14$, $p = .889$. Therefore, $Ho3_1$ was retained. The percentage of students scoring in the basic category in ability grouped classrooms ($M = 31.06$, $SD = 11.38$) was similar to those who were in a classroom not grouped by ability ($M = 31.55$, $SD = 12.78$). The 95% confidence interval for the difference in means was -7.51 to 6.53. The η^2 index was < .001, which indicated a small effect size. Figure 10 shows the distributions for the two groups.

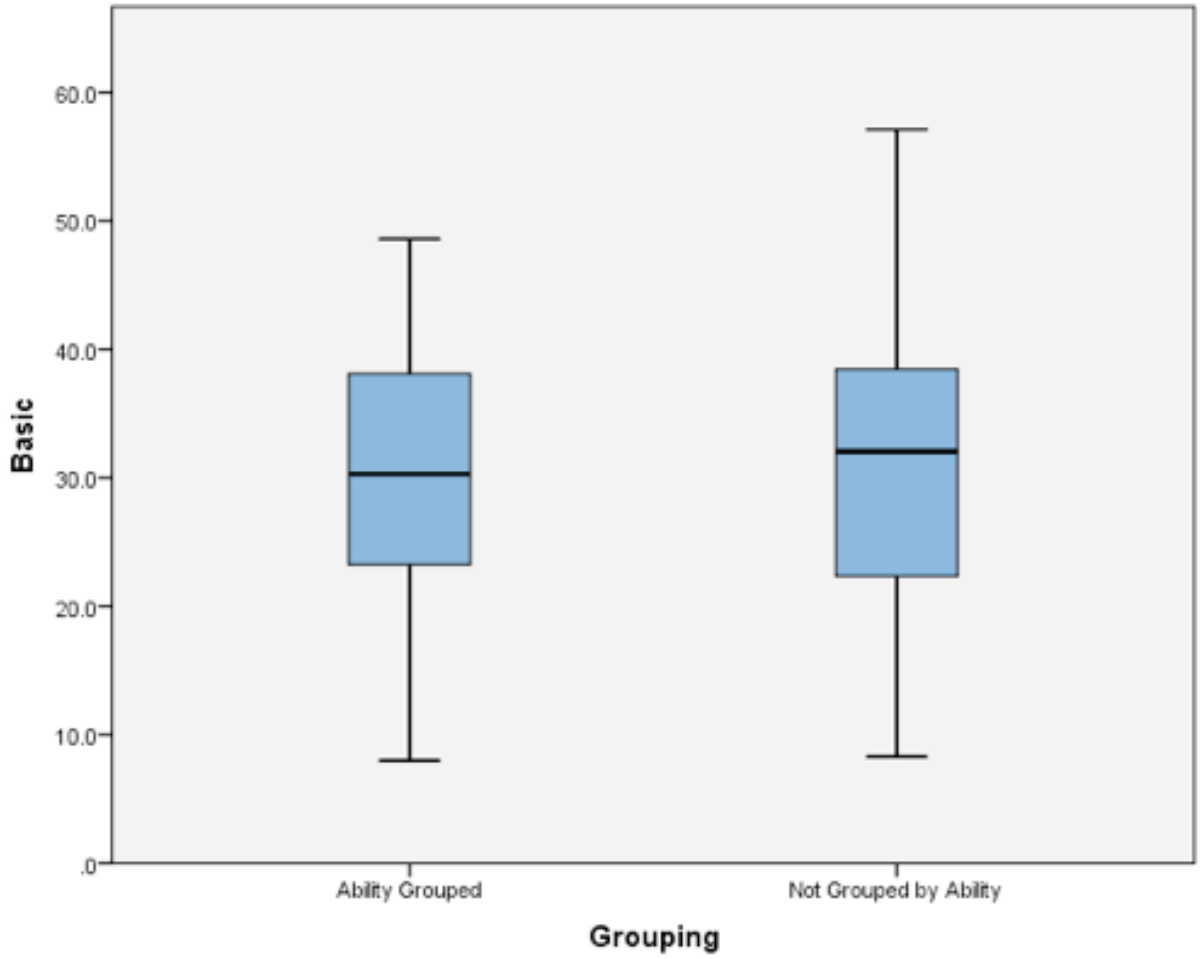

Figure 10: 5th Grade Math: Percentage of Students Scoring in the Basic Proficiency Level

H_o3_3: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the proficient level for fifth grade math between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in math of fifth grade students scoring in the proficient category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(50) = .23$, $p = .817$. Therefore, H_o3_3 was retained. The percentage of students scoring in the proficient category in ability grouped classrooms ($M = 32.41$, $SD = 10.96$) was similar to those who were in a classroom not grouped by ability ($M = 20.64$, $SD = 9.86$). The 95% confidence interval for the difference in means was -5.21 to 6.58. The η^2 index was .001, which indicated a small effect size. Figure 11 shows the distributions for the two groups.

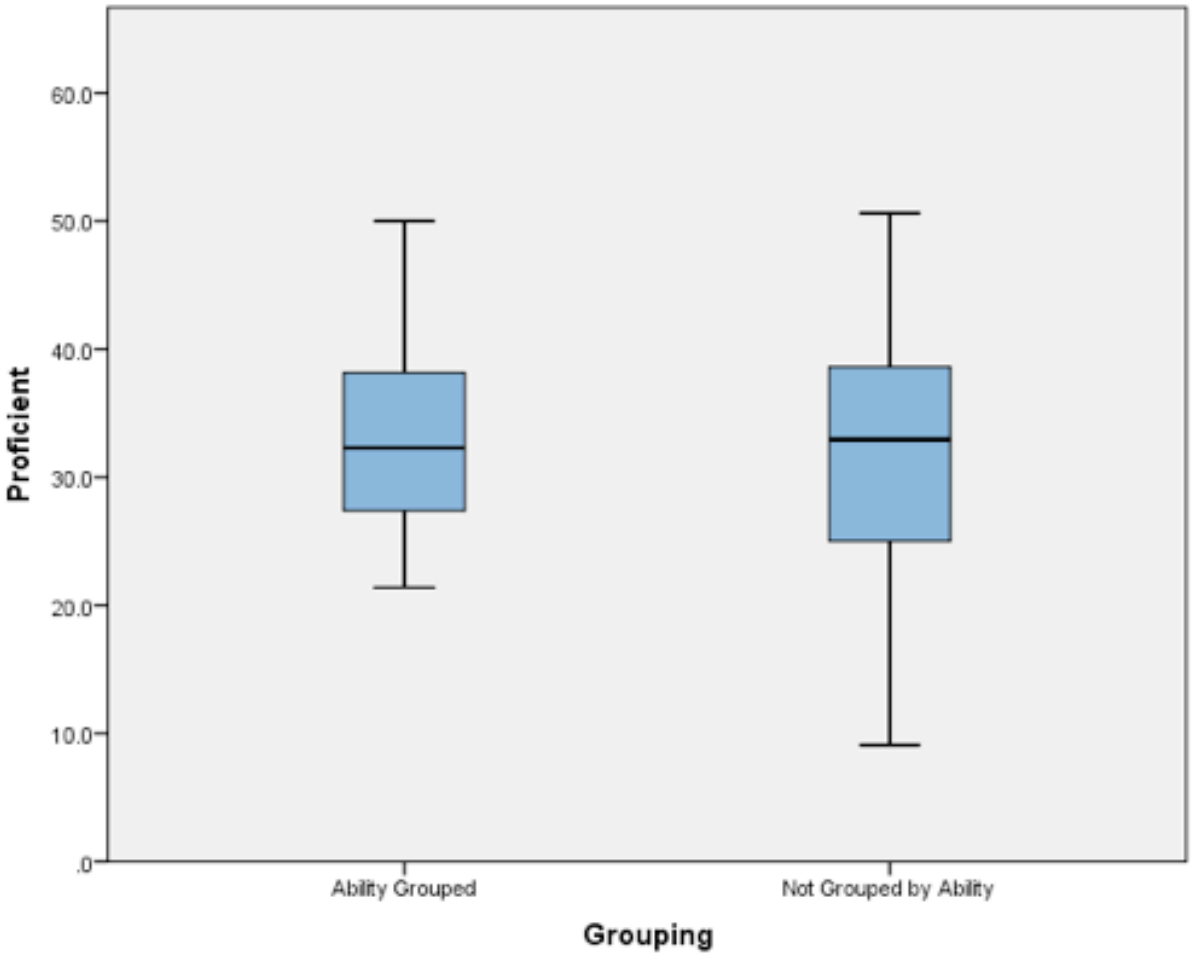

Figure 11: 5th Grade Math: Percentage of Students Scoring in the Proficient Proficiency Level

H_o4_4: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the advanced level for fifth grade math between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in math of fifth grade students scoring in the advanced category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped

by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(50) = .50$, $p = .619$. Therefore, H_o3_4 was retained. The percentage of students scoring in the advanced category in ability grouped classrooms ($M = 20.64$, $SD = 14.75$) was similar to those who were in a classroom not grouped by ability ($M = 22.85$, $SD = 15.99$). The 95% confidence interval for the difference in means was -11.11 to 6.68. The η^2 index was .005, which indicated a small effect size. Figure 12 shows the distributions for the two groups.

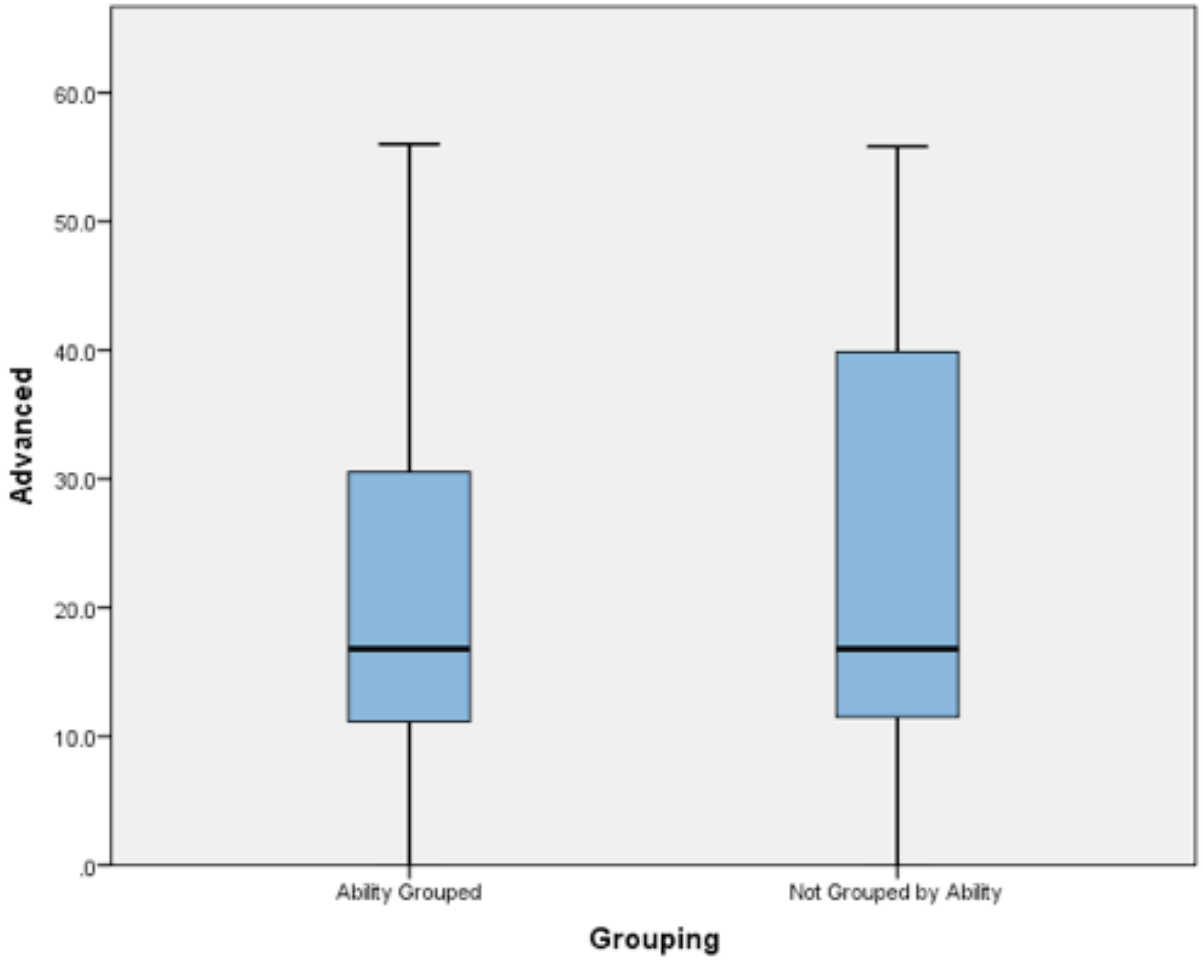

Figure 12: 5th Grade Math: Percentage of Students Scoring in the Advanced Proficiency Level

Research Question 4

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for third grade reading between schools that have students grouped by ability and schools that do not group students by ability?

H_o4_1: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the below basic level for third grade reading between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in reading of third grade students scoring in the below basic category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(54) = .67$, $p = .505$. Therefore, H_o4_1 was retained. The percentage of students scoring in the below basic category in ability grouped classrooms ($M = 12.39$, $SD = 11.19$) was similar to those who were in a classroom not grouped by ability ($M = 10.66$, $SD = 7.57$). The 95% confidence interval for the difference in means was -3.44 to 6.91. The η^2 index was .009, which indicated a small effect size. Figure 13 shows the distributions for the two groups.

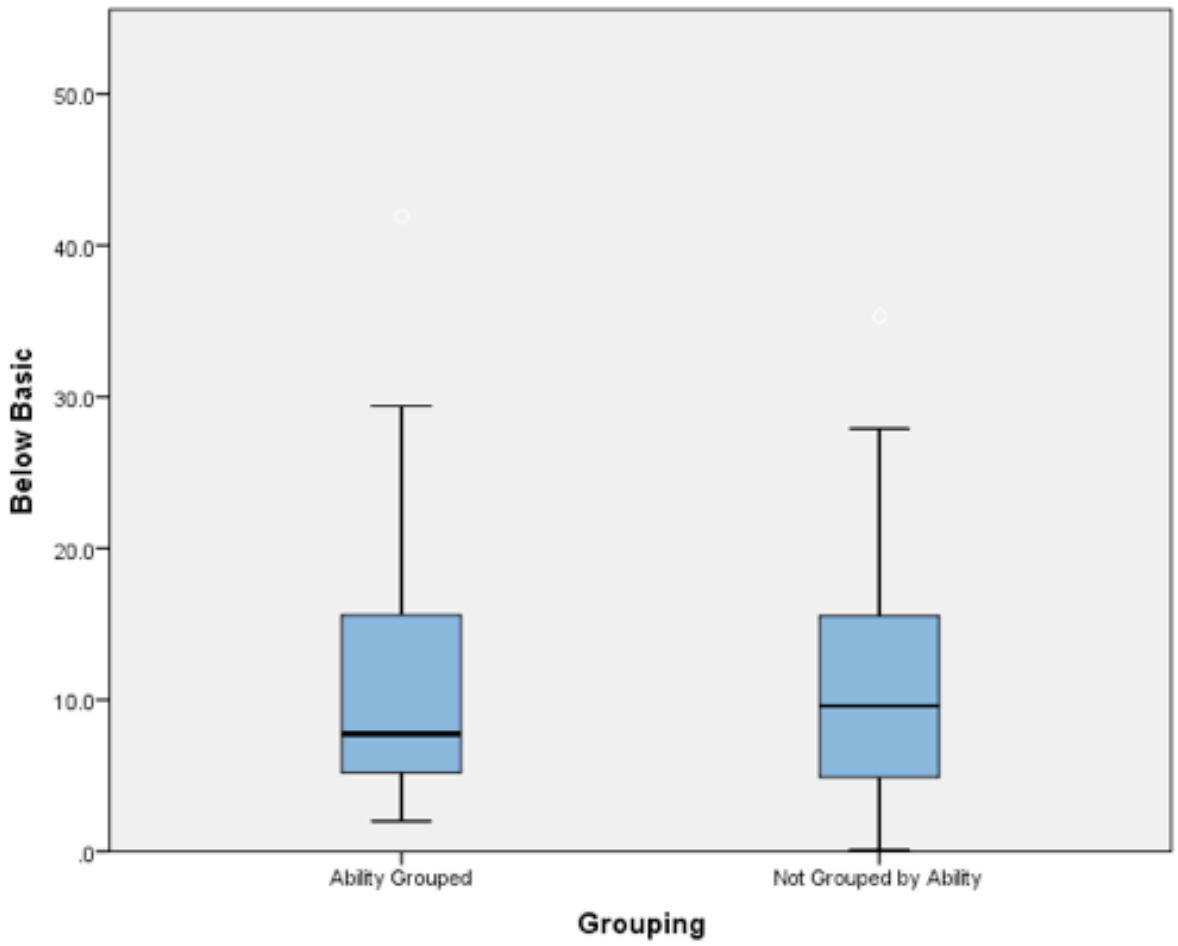

Figure 13: 3rd Grade Reading: Percentage of Students Scoring in the Below Basic Proficiency Level

H_o4_2: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the basic level for third grade reading between schools that have students grouped by ability and schools that do not group students by ability

An independent samples t-test was conducted to evaluate whether the mean percentages in reading of third grade students scoring in the basic category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped by ability. The percentage of students scoring in the category was the test variable and the

grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(54) = 1.15$, $p = .136$. Therefore, H_o4_2 was retained. The percentage of students scoring in the basic category in ability grouped classrooms ($M = 38.47$, $SD = 9.24$) was similar to those who were in a classroom not grouped by ability ($M = 43.18$, $SD = 10.98$). The 95% confidence interval for the difference in means was -10.96 to 1.15. The η^2 index was .024, which indicated a small effect size. Figure 14 shows the distributions for the two groups.

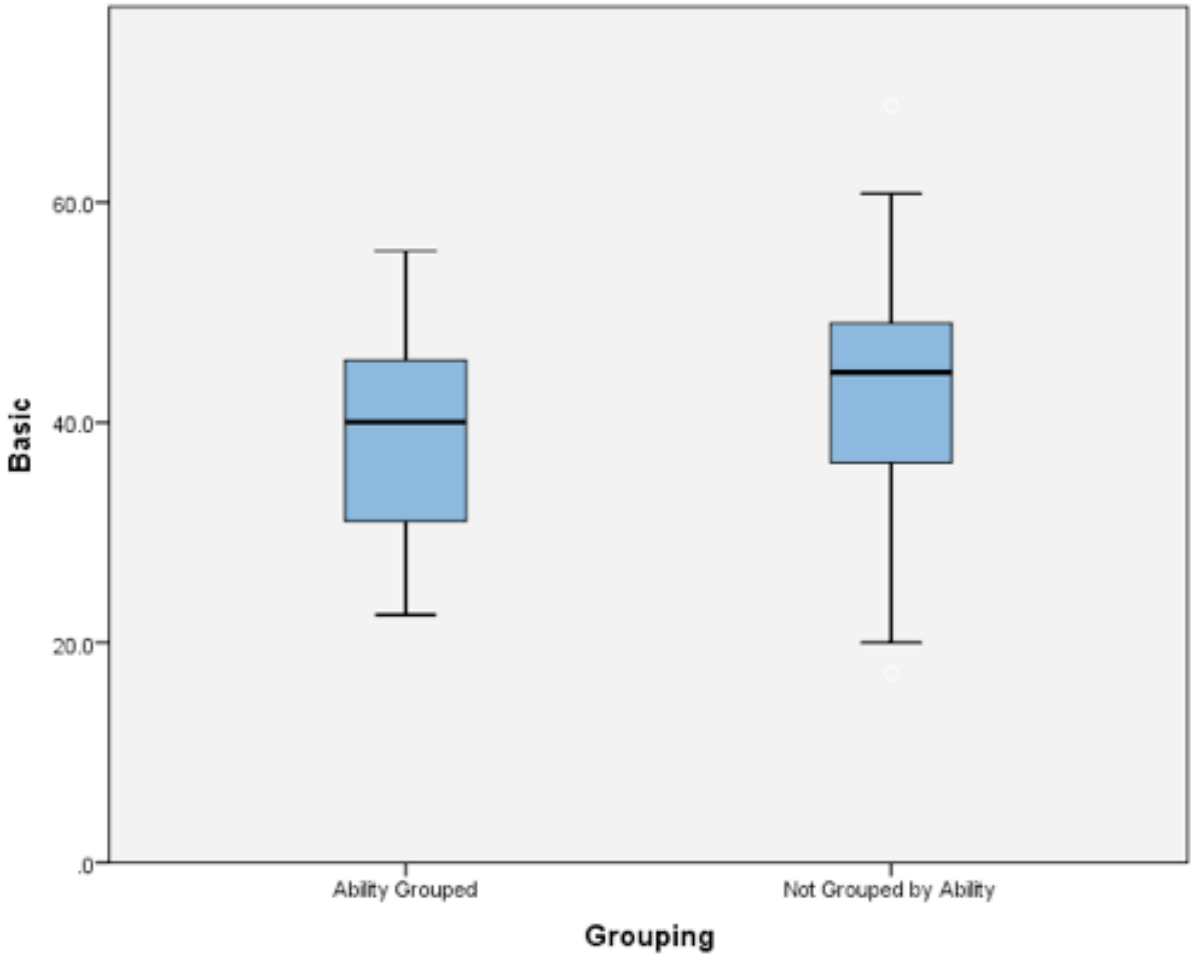

Figure 14: 3rd Grade Reading: Percentage of Students Scoring in the Basic Proficiency Level

H_o4_3: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the proficient level for third grade

reading between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in reading of third grade students scoring in the proficient category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(54) = 1.23$, $p = .223$. Therefore, H_o4_3 was retained. The percentage of students scoring in the proficient category in ability grouped classrooms ($M = 38.38$, $SD = 9.10$) was similar to those who were in a classroom not grouped by ability ($M = 34.73$, $SD = 10.33$). The 95% confidence interval for the difference in means was -2.28 to 9.58. The η^2 index was .027, which indicated a small to medium effect size. Figure 15 shows the distributions for the two groups.

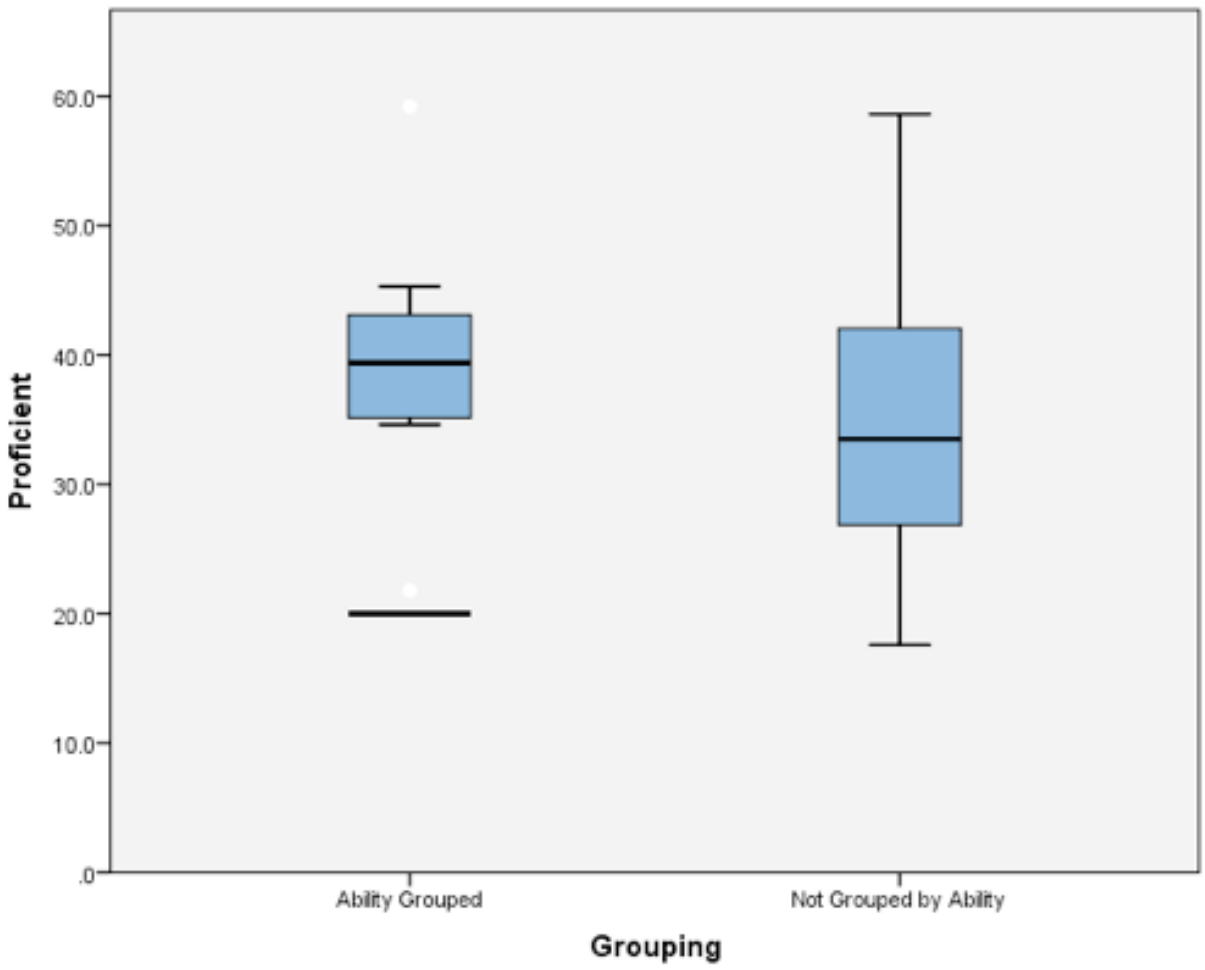

Figure 15: 3rd Grade Reading: Percentage of Students Scoring in the Proficient Proficiency Level

H_o4_4: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the advanced level for third grade reading between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in reading of third grade students scoring in the advanced category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped by ability. The percentage of students scoring in the category was the test variable and

the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(54) = .31$, $p = .756$. Therefore, H_o4_4 was retained. The percentage of students scoring in the advanced category in ability grouped classrooms ($M = 10.76$, $SD = 8.50$) was similar to those who were in a classroom not grouped by ability ($M = 11.43$, $SD = 6.70$). The 95% confidence interval for the difference in means was -4.97 to 3.63. The η^2 index was .002, which indicated a small effect size. Figure 16 shows the distributions for the two groups.

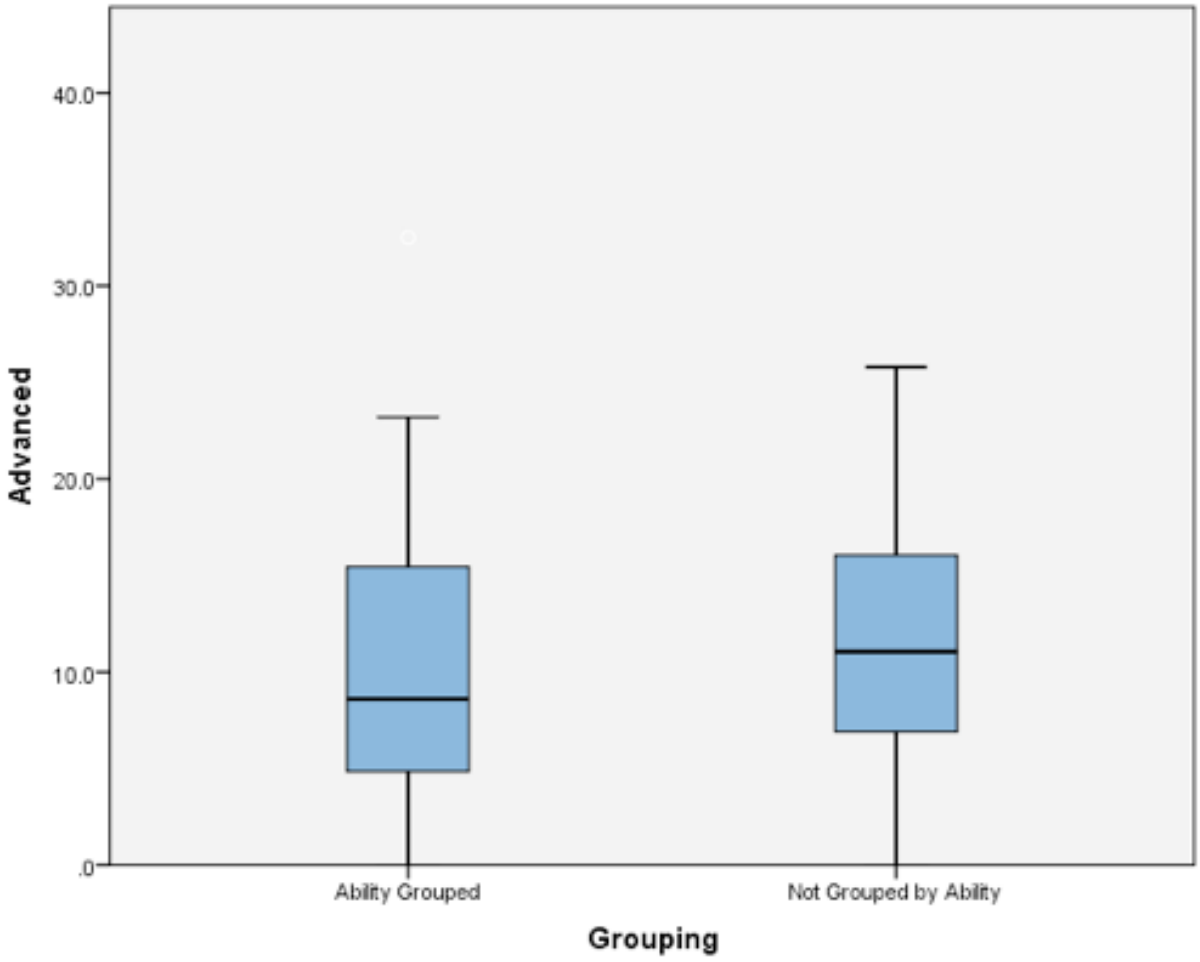

Figure 16: 3rd Grade Reading: Percentage of Students Scoring in the Advanced Proficiency Level

Research Question 5

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for fourth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

H_o5_1: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the below basic level for fourth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in reading of fourth grade students scoring in the below basic category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(53) = .57$, $p = .573$. Therefore, H_o5_1 was retained. The percentage of students scoring in the below basic category in ability grouped classrooms ($M = 12.88$, $SD = 10.44$) was similar to those who were in a classroom not grouped by ability ($M = 11.50$, $SD = 7.51$). The 95% confidence interval for the difference in means was -6.37 to 3.91. The η^2 index was .006, which indicated a small effect size. Figure 17 shows the distributions for the two groups.

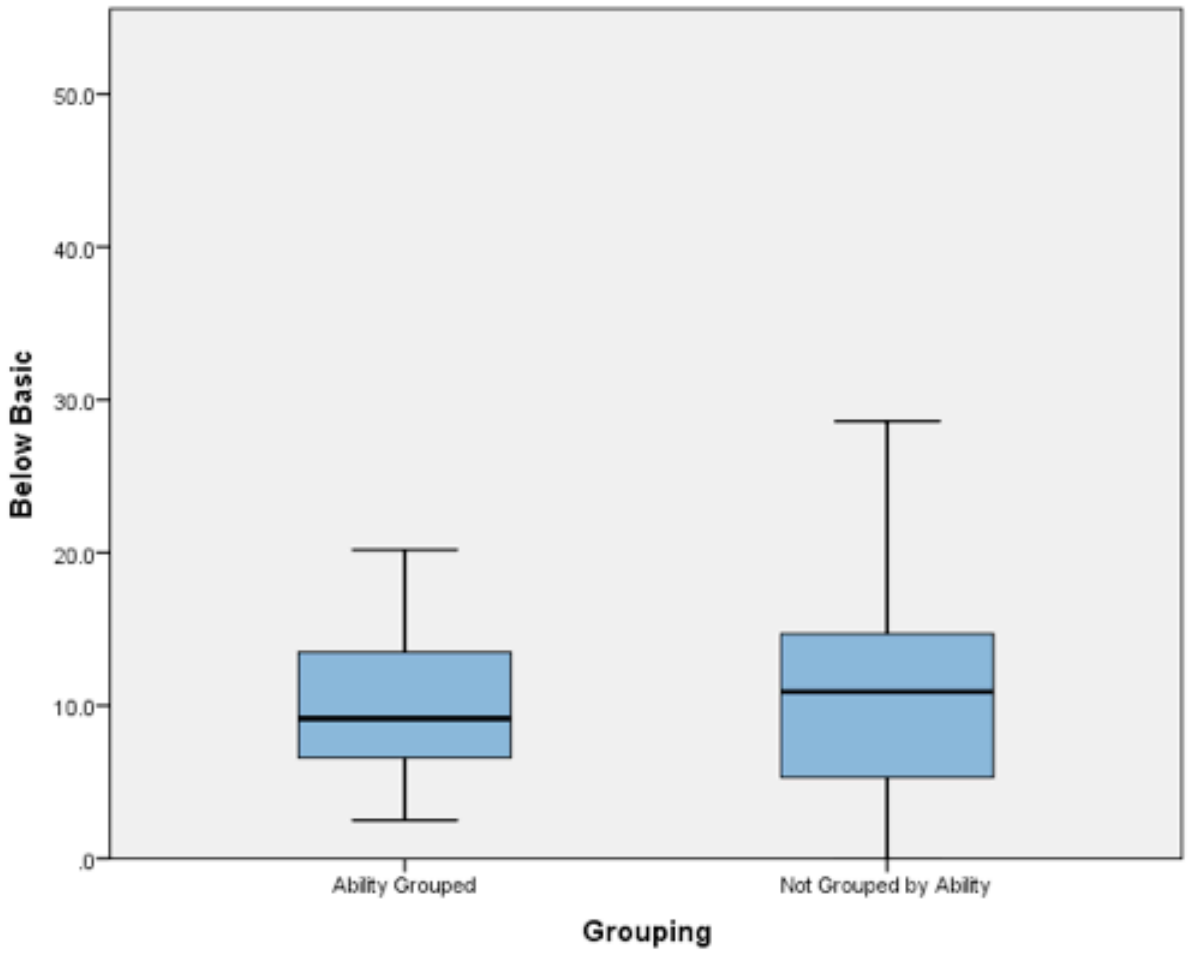

Figure 17: 4th Grade Reading: Percentage of Students Scoring in the Below Basic Proficiency Level

H_o5_2: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the basic level for fourth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in reading of fourth grade students scoring in the basic category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped

by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(53) = .31$, $p = .759$. Therefore, H_o5_2 was retained. The percentage of students scoring in the basic category in ability grouped classrooms ($M = 43.32$, $SD = 9.42$) was similar to those who were in a classroom not grouped by ability ($M = 42.48$, $SD = 9.95$). The 95% confidence interval for the difference in means was -4.65 to 6.33. The η^2 index was .002, which indicated a small effect size. Figure 18 shows the distributions for the two groups.

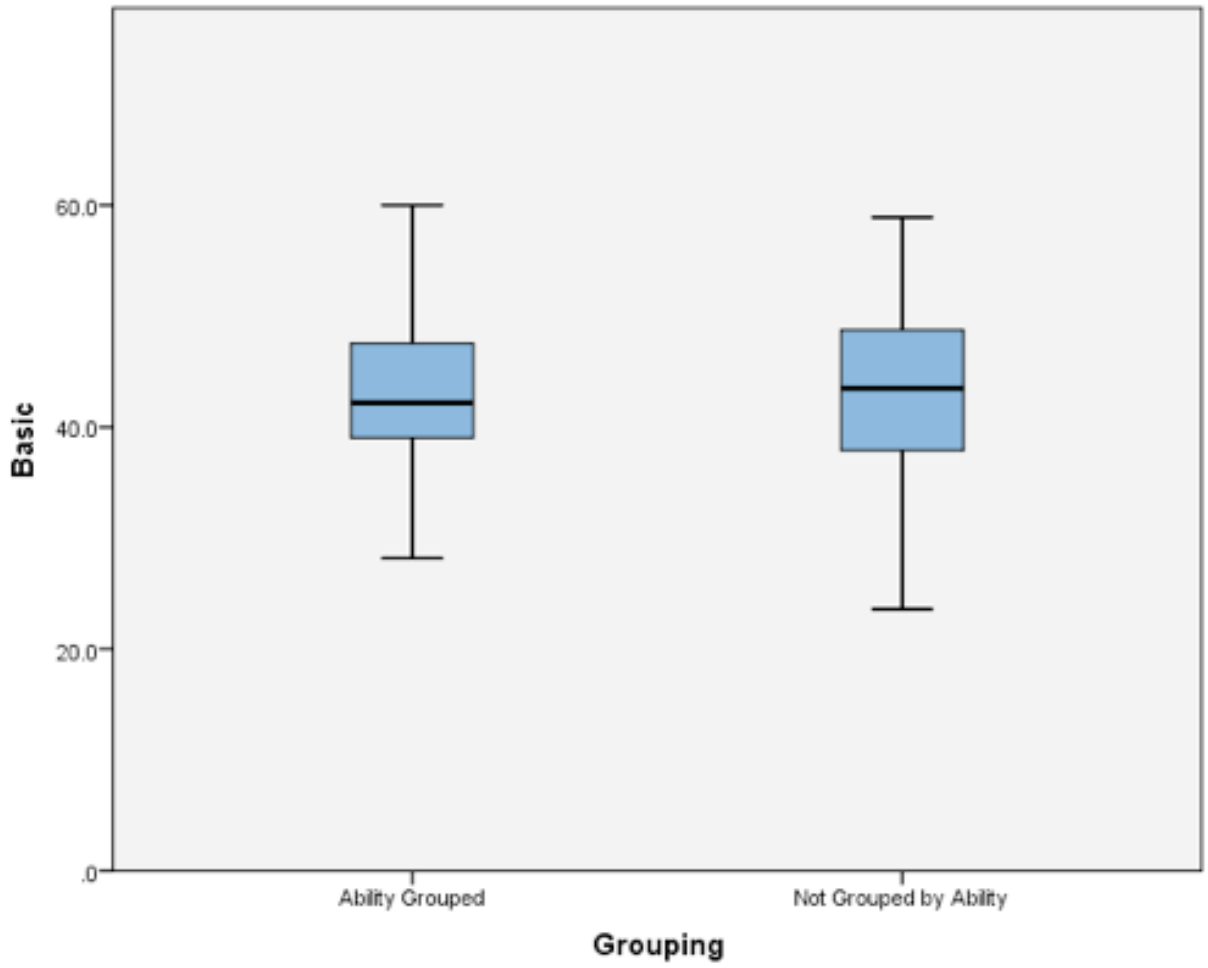

Figure 18: 4th Grade Reading: Percentage of Students Scoring in the Basic Proficiency Level

H_o5_3: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the proficient level for fourth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in reading of fourth grade students scoring in the proficient category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(53) = .48$, $p = .634$. Therefore, H_o5_3 was retained. The percentage of students scoring in the proficient category in ability grouped classrooms ($M = 32.27$, $SD = 9.84$) was similar to those who were in a classroom not grouped by ability ($M = 33.49$, $SD = 8.73$). The 95% confidence interval for the difference in means was -6.37 to 3.91. The η^2 index was .004, which indicated a small effect size. Figure 19 shows the distributions for the two groups.

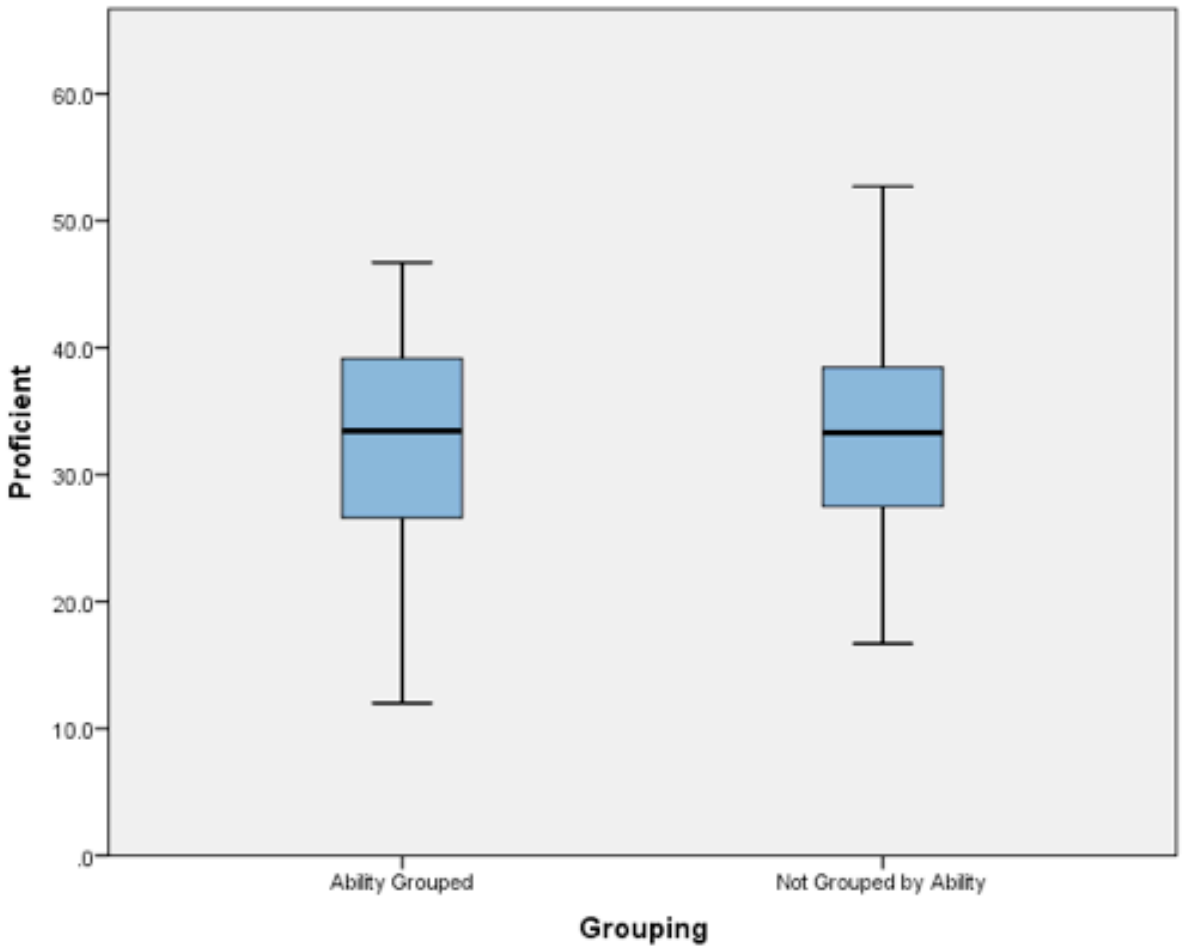

Figure 19: 4th Grade Reading: Percentage of Students Scoring in the Proficient Proficiency Level

H_o5_4: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the advanced level for fourth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in reading of fourth grade students scoring in the advanced category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not

grouped by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(53) = .51$, $p = .610$. Therefore, H_o5_4 was retained. The percentage of students scoring in the advanced category in ability grouped classrooms ($M = 11.54$, $SD = 6.11$) was similar to those who were in a classroom not grouped by ability ($M = 12.53$, $SD = 7.32$). The 95% confidence interval for the difference in means was -4.88 to 2.89. The η^2 index was .005, which indicated a small effect size. Figure 20 shows the distributions for the two groups.

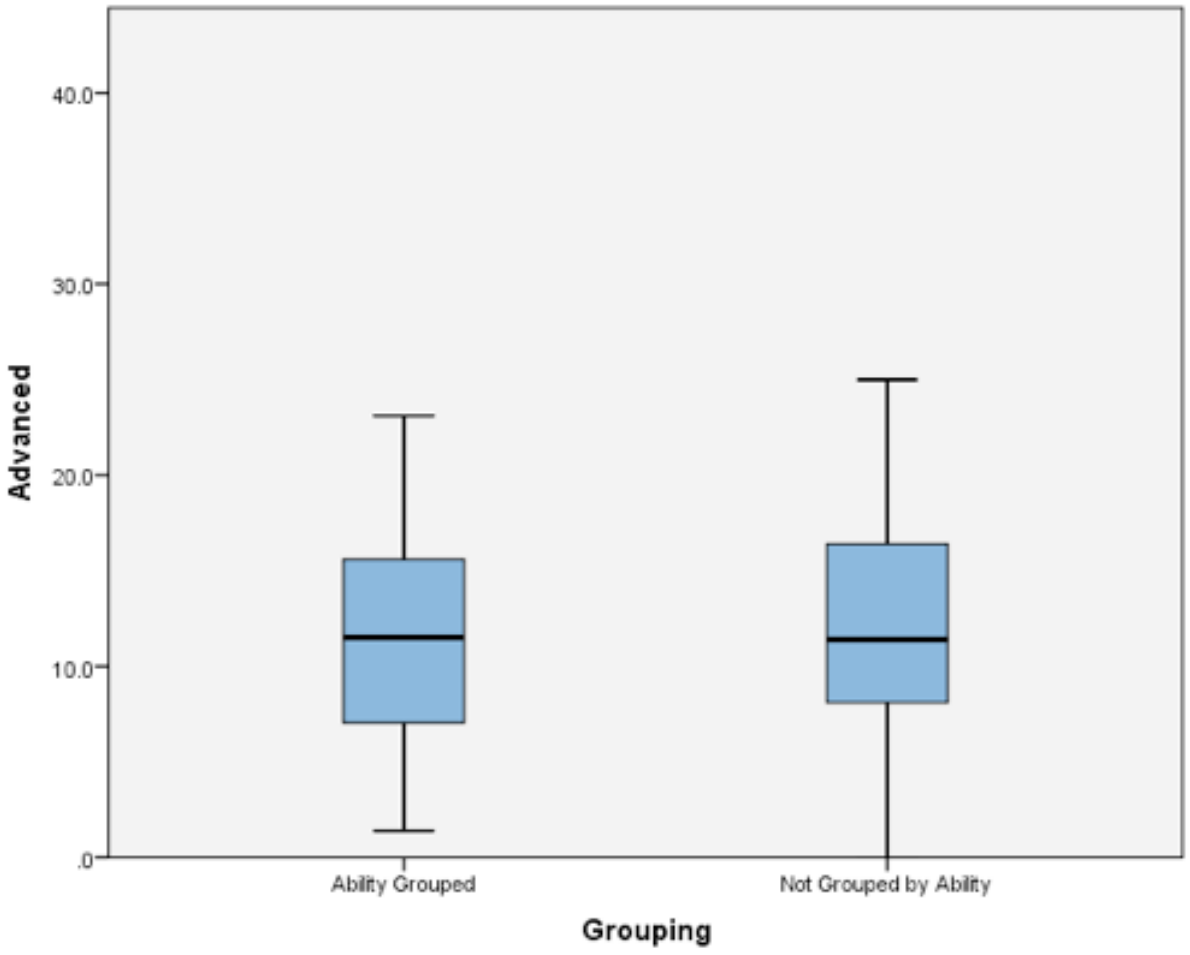

Figure 20: 4th Grade Reading: Percentage of Students Scoring in the Advanced Proficiency Level

Research Question 6

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for fifth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

H_o6_1: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the below basic level for fifth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in reading of fifth grade students scoring in the below basic category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(49) = 1.47$, $p = .149$. Therefore, H_o6_1 was retained. The percentage of students scoring in the below basic category in ability grouped classrooms ($M = 7.50$, $SD = 6.52$) was similar to those who were in a classroom not grouped by ability ($M = 10.60$, $SD = 7.72$). The 95% confidence interval for the difference in means was -7.35 to 1.15. The η^2 index was .042, which indicated a small to medium effect size. Figure 21 shows the distributions for the two groups.

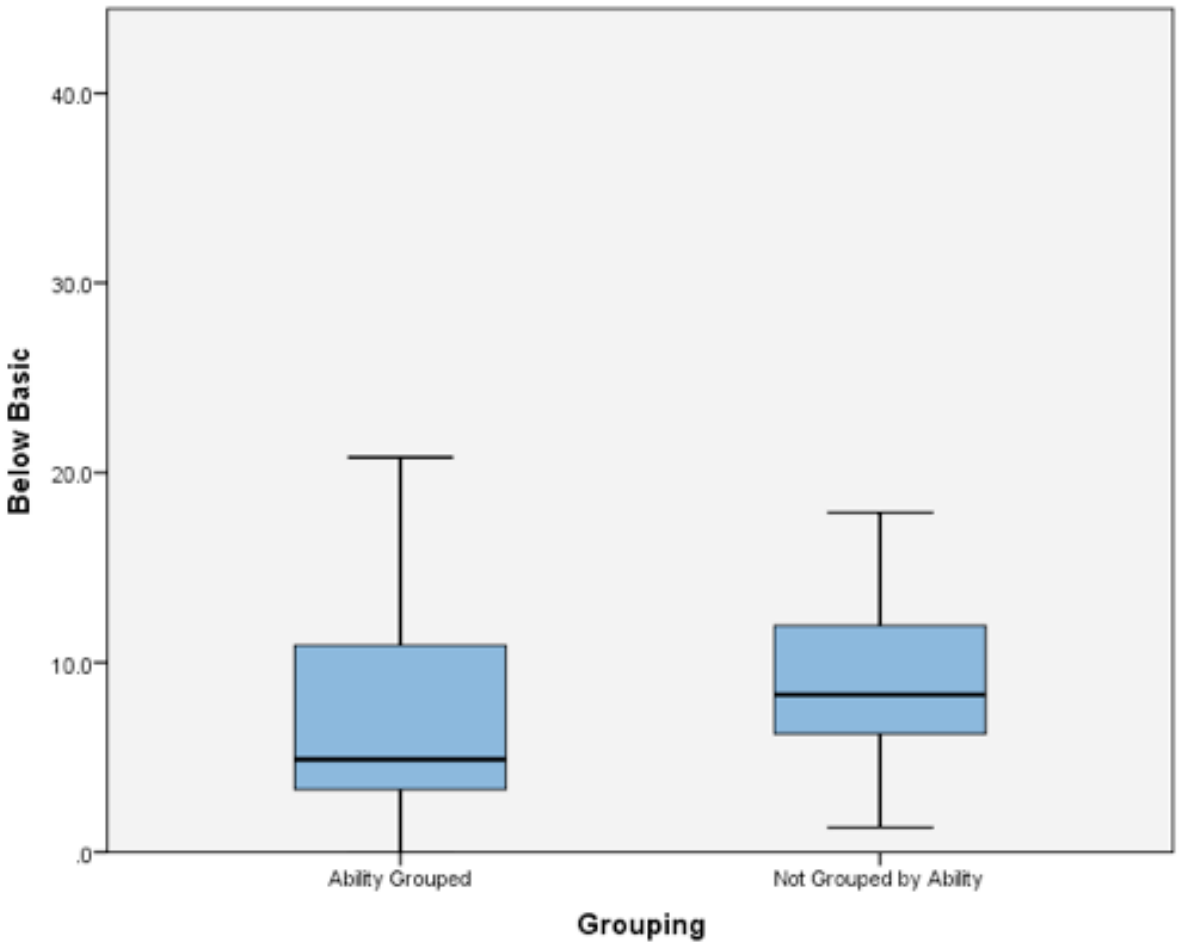

Figure 21: 5th Grade Reading: Percentage of Students Scoring in the Below Basic Proficiency Level

H_o6_2: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the basic level for fifth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in reading of fifth grade students scoring in the basic category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped by ability. The percentage of students scoring in the category was the test variable and the

grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(49) = 1.16$, $p = .254$. Therefore, H_o6_2 was retained. The percentage of students scoring in the basic category in ability grouped classrooms ($M = 41.11$, $SD = 10.09$) was similar to those who were in a classroom not grouped by ability ($M = 37.73$, $SD = 10.14$). The 95% confidence interval for the difference in means was -2.51 to 9.28. The η^2 index was .027, which indicated a small effect size. Figure 21 shows the distributions for the two groups.

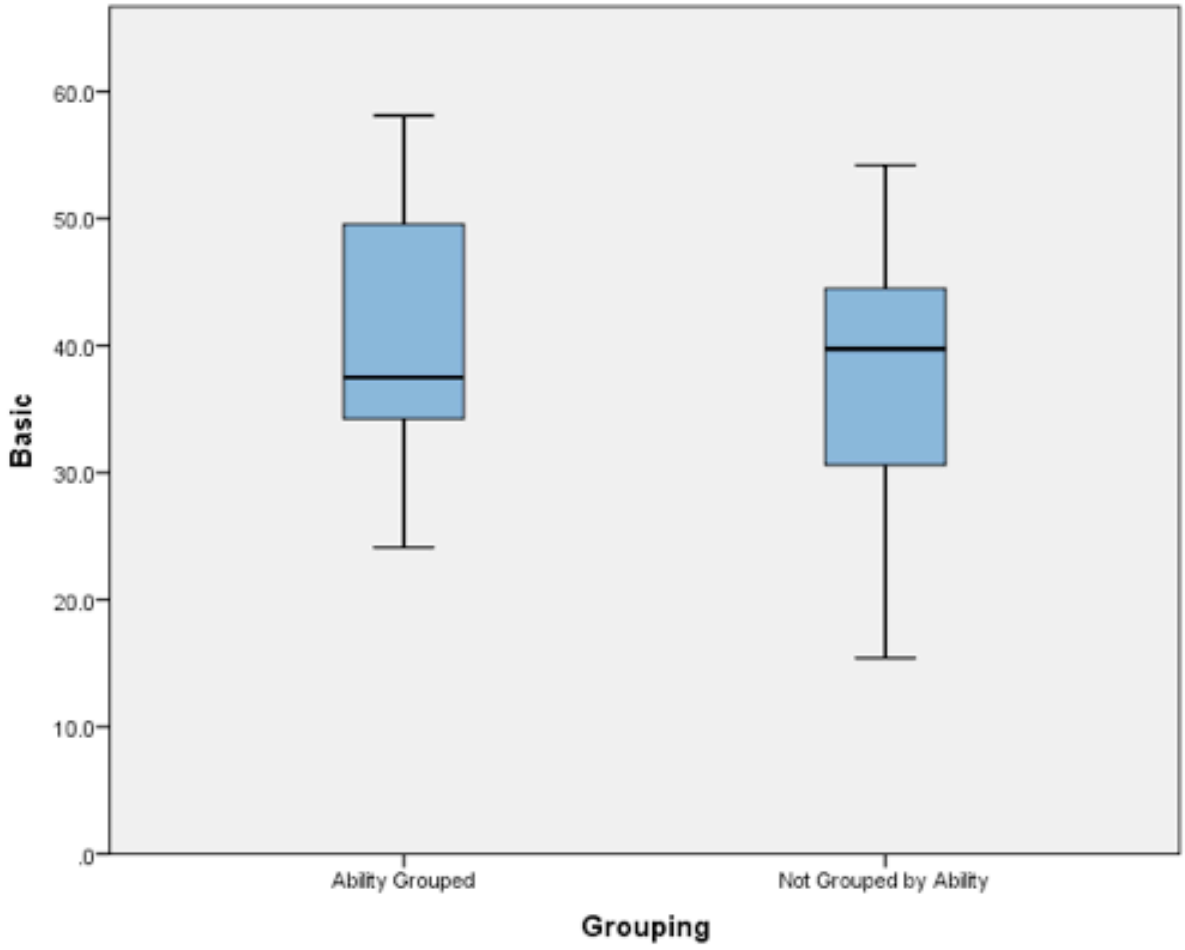

Figure 22: 5th Grade Reading: Percentage of Students Scoring in the Basic Proficiency Level

H_o6_3: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the proficient level for fifth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in reading of fifth grade students scoring in the proficient category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not grouped by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(49) = .19$, $p = .847$. Therefore, H_o6_3 was retained. The percentage of students scoring in the proficient category in ability grouped classrooms ($M = 43.04$, $SD = 9.94$) was similar to those who were in a classroom not grouped by ability ($M = 43.62$, $SD = 10.49$). The 95% confidence interval for the difference in means was -6.56 to 5.40. The η^2 index was .001, which indicated a small effect size. Figure 23 shows the distributions for the two groups.

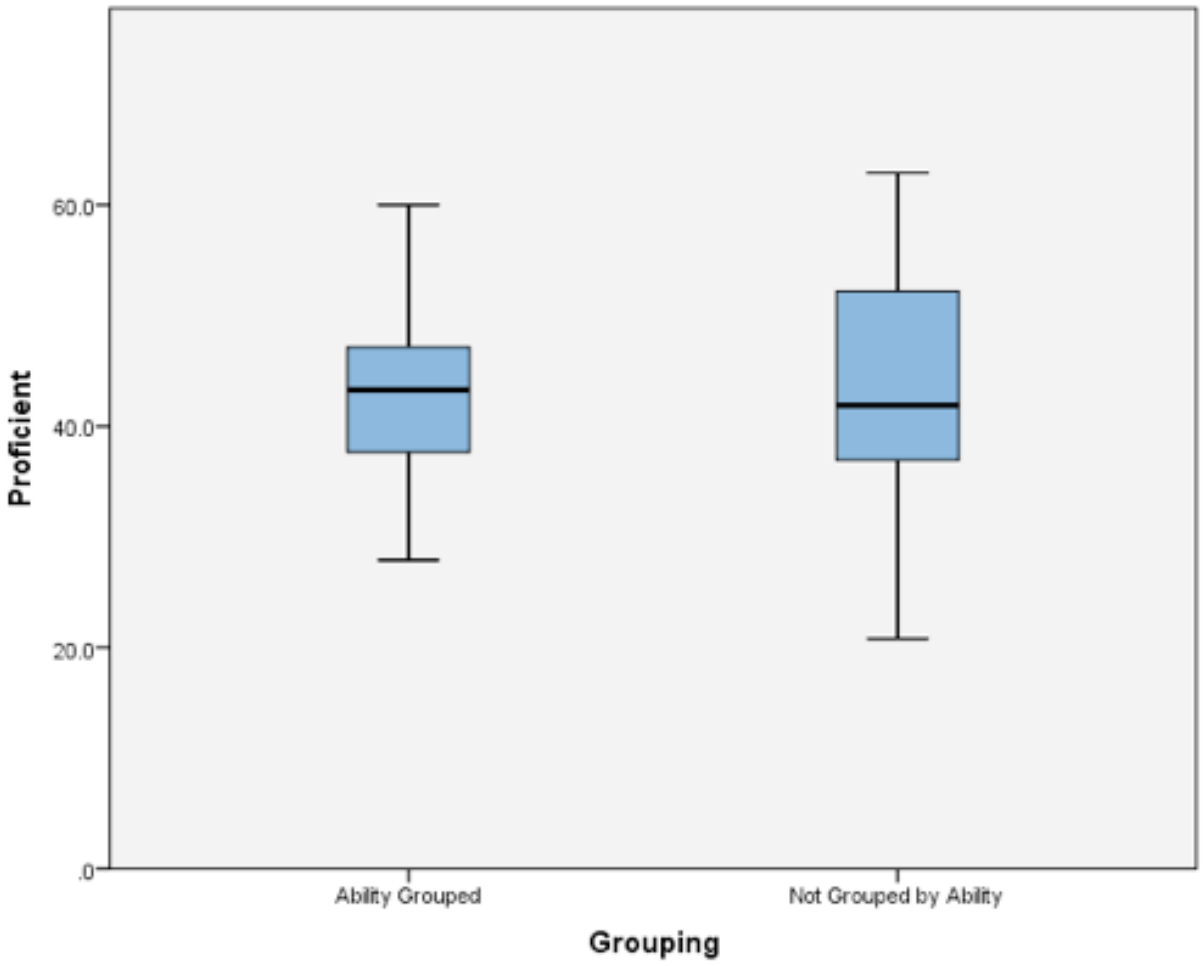

Figure 23: 5th Grade Reading: Percentage of Students Scoring in the Proficient Proficiency Level

H_o6_4: There is no significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) in the advanced level for fifth grade reading between schools that have students grouped by ability and schools that do not group students by ability?

An independent samples t-test was conducted to evaluate whether the mean percentages in reading of fifth grade students scoring in the advanced category differed between those receiving instruction in an ability grouped setting and those who were placed in a classroom not

grouped by ability. The percentage of students scoring in the category was the test variable and the grouping variable was the type of grouping: ability grouped or not grouped by ability. The test was not significant, $t(49) = .16$, $p = .871$. Therefore, H_o6_4 was retained. The percentage of students scoring in the advanced category in ability grouped classrooms ($M = 8.35$, $SD = 5.18$) was similar to those who were in a classroom not grouped by ability ($M = 8.06$, $SD = 6.61$). The 95% confidence interval for the difference in means was -3.27 to 3.85. The η^2 index was $< .001$, which indicated a small effect size. Figure 24 shows the distributions for the two groups.

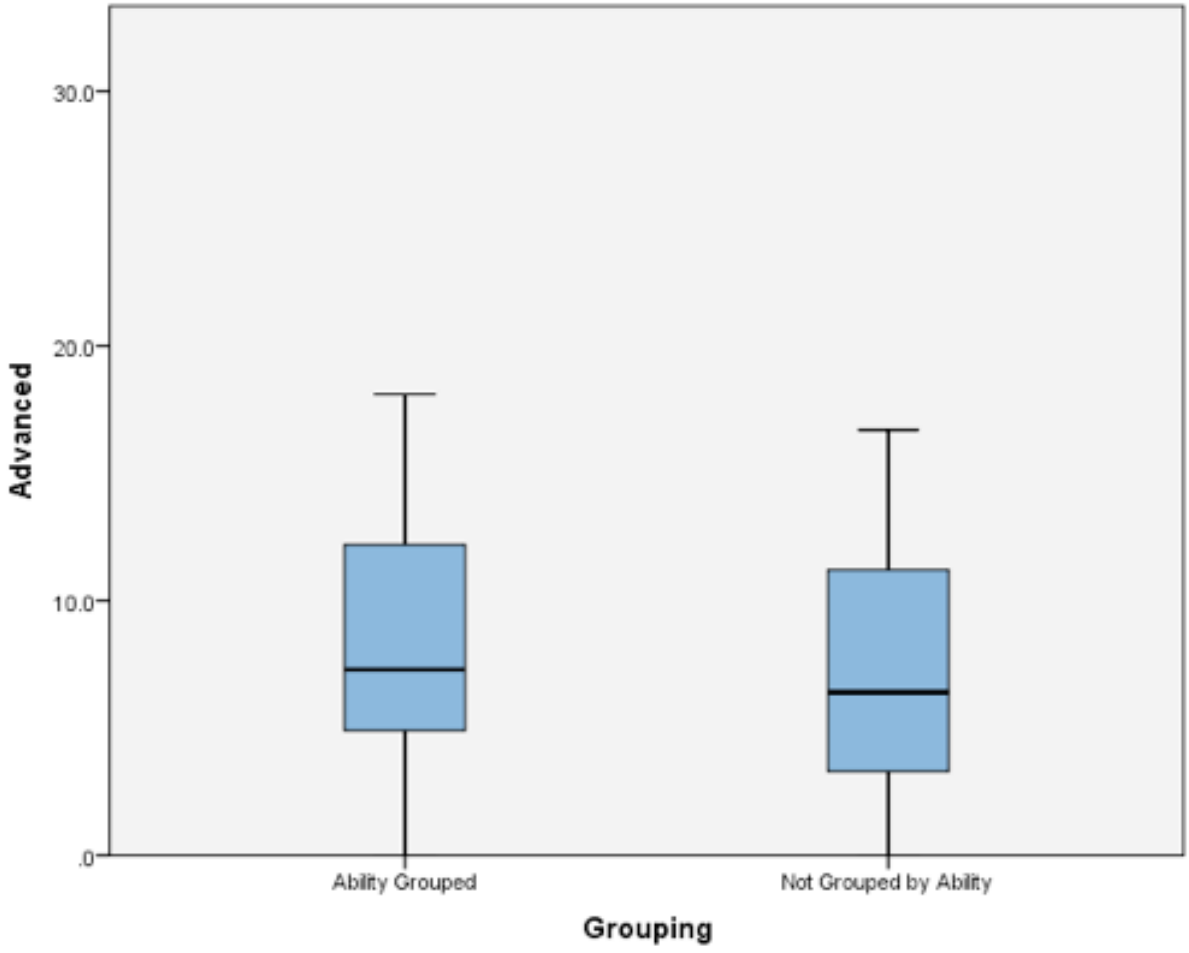

Figure 24: 5th Grade Reading: Percentage of Students Scoring in the Advanced Proficiency Level

Summary

In this chapter percentages of students scoring in each reporting category (below basic, basic, proficient, and advanced) on the 2012-2013 TCAP were obtained for grades 3, 4, and 5. School administrators were contacted and the data were sorted based on the type of grouping implemented: ability grouping or not grouped by ability. Approximately fifty-five schools were contacted. Of those 55, approximately 20 grouped by ability and 30 did not group by ability. There was variation in the number of schools that grouped by ability and did not group by ability due to differences in school make-up (some schools were only grades 4 and 5) and some school only grouped in certain grade levels. There were six research questions and 24 null hypotheses. The data were obtained from the Tennessee Department of Education website and individual school administrators.

A series of independent t-tests were run to analyze if there was a significant difference in percentage of students in each reporting category based on the type of grouping implemented. The results showed that there is not a significant different in percentages of students in all categories based on the type of grouping. Therefore, all null hypotheses were retained. Chapter 5 is a discussion of the conclusions of this study, how they are related to other literature, and suggestions for further research.

CHAPTER 5

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Grouping students based on ability has been a common practice in education in order to meet the individual needs of students. Research has provided evidence for both the positive and negative aspects of ability grouping.

Summary of Results

The purpose of this study was to evaluate if there is a significant relationship between the percentage of students scoring in each proficiency level (below basic, basic, proficient, and advanced) on the 2012-2013 TCAP assessment in grades 3, 4, and 5 in both reading and math between schools that group students by ability and those that do not group by ability. TCAP data were collected from the Tennessee Department of Education website. Individual school administrators were contacted to determine if schools grouped students by ability for reading and math in grades 3, 4, and 5 or if they do not group by ability. Table 1 shows the number of schools that grouped by ability or did not group by ability in both subjects and in each grade level.

Table 1

Schools Reporting Ability Grouped or Not Grouped By Ability

		Grouped by Ability	Not Grouped by Ability
3rd Grade	Reading	16	40
	Math	16	40
4th Grade	Reading	20	35
	Math	21	34
5th Grade	Reading	19	32
	Math	20	32

The table shows the number of schools that reported either grouping by ability or not grouping by ability in the 2012-2013 school year by grade and subject.

Research Question 1

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for third grade math between schools that have students grouped by ability and schools that have students not grouped by ability?

An independent samples t-test was conducted to evaluate if there is a significant difference in the percentage of students in each of the proficiency levels for third grade math between schools that group by ability and those that do not group by ability. The test results showed that there is no significant difference between students receiving instruction in ability grouped classrooms and those who receive instruction in classrooms not grouped by ability in all four proficiency levels.

Research Question 2

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for fourth grade math between schools that have students grouped by ability and schools that have students not grouped by ability?

An independent samples t-test was conducted to determine if there is a statistically significant difference between the percentages of students that score in each of the four proficiency levels in fourth grade math between students receiving instruction in ability grouped classrooms and those who are not grouped by ability. The results of the test indicated that there is no significance between students that receive instruction in ability grouped classrooms and those that do not group by ability in fourth grade math.

Research Question 3

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for fifth grade math between schools that have students grouped by ability and schools that have students not grouped by ability?

An independent samples t-test was conducted to analyze if there is a significant difference in the percentage of students scoring in each of the TCAP proficiency levels in fifth grade math between students receiving instruction in classrooms that are grouped by ability and those that are not grouped by ability. The results of the test revealed that there is no significant difference. Therefore, there is no difference in the percentages of students scoring in each of the proficiency levels in fifth grade math between students receiving instruction in classrooms that group by ability and those that do not group by ability.

Research Question 4

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for third grade reading between schools that have students grouped by ability and schools that have students not grouped by ability?

An independent samples t-test was run to determine if there is a significant different in the percentage of students scoring in each of the proficiency levels in third grade reading on the TCAP test between students receiving instruction in classrooms grouped by ability and those not grouped by ability. The test results revealed that there is not a significant difference between students who are grouped by ability for instruction and those that are not grouped by ability in third grade reading.

Research Question 5

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for fourth grade reading between schools that have students grouped by ability and schools that have students not grouped by ability?

An independent samples t-test was conducted to determine if there was a significant difference in each proficiency level for fourth grade reading between schools that grouped by ability and those that did not group by ability. The test results indicated that there is not a significant difference in fourth grade reading between students who received instruction in classrooms grouped by ability and those that were not grouped by ability.

Research Question 6

Is there a significant difference in the percentage of students in each of the Tennessee Comprehensive Achievement Program (TCAP) proficiency levels (Advanced, Proficient, Basic, Below Basic) for fifth grade reading between schools that have students grouped by ability and schools that have students not grouped by ability?

An independent samples t-test was conducted to determine if there is a significant difference in the percentage of students scoring in each TCAP proficiency level in fifth grade reading between students receiving instruction in classrooms grouped by ability and those who were not grouped by ability. The results showed that there was not a significant difference between students receiving instruction in classrooms grouped by ability and those not grouped by ability in fifth grade reading for each of the four proficiency levels.

Implications for Practice

Research in the literature uncovered mixed results as to the impact that grouping by ability and not grouping by ability have on student achievement (Reuman, 1989). Overall, researchers concluded that students who are high achievers typically benefit from instruction in ability groups, whereas students who are middle or low performers do not receive any academic benefits (Plunkett & Kronborg, 2011). As it is the responsibility of the educator to prepare the whole child for the future, it is important to balance the academic impact of an instructional strategy with the social-emotional effects. Researchers discovered that ability grouping in post-primary grades does not tend to have a significant impact on student self-concept and perceptions of learning; however, researchers have found that students in primary grades tend to have a lower self-concept scores as well as perception of school (Hong et al, 2012; Ireson et al., 2001). There is a positive correlation between student self-concept and student achievement.

Based on the conflicting research surrounding ability grouping and the academic impact of this instructional practice, further research is necessary. In examining the results of this study, there is no significant difference between students receiving instruction in math and reading in an ability grouped setting and those receiving instruction not in an ability grouped setting based on the percentage of students scoring in each TCAP proficiency level in grades 3, 4, and 5. The results of this study should be examined with the limitations in mind. Betts (2011) stated that much of the discrepancies in research regarding ability grouping is based on miscommunication or a misunderstanding of definitions or poor measurements.

The results of this study are similar to the findings of Slavin (1990) in his meta-analysis of literature regarding ability grouping. He found that when converging the results of over 29 international studies, there was no correlation between ability grouping and academic achievement. Reuman (1989) and Wouters et al. (2012) similarly found that overall students do not score significantly higher when grouped by ability; Reuman discovered that students actually scored higher on a sixth grade math assessment from the beginning of the year to the end of the year when they were placed in a classroom that was not grouped by ability.

While some research supports academic benefits of ability grouping (Duflo et al., 2009), most previous research took into consideration other factors such as socioeconomic status, Title I schools, class size, or individual intervention or acceleration programs. Some research studies indicated students performing above grade level show an increase in growth if they participate in a specific gifted and talents program; however, many gifted and talented programs are only implemented for a portion of the day and students spend other portions of the day in a mixed ability grouped setting (Castambis & Buttaro, 2012). The practice of placing students in a mixed ability grouped setting is to increase socialization with students who have diverse learning styles

and abilities and to provide students an opportunity to learn how to work with others as in a real-world setting.

The results of this study were based on the type of grouping implemented and TCAP achievement scores. However, there is something to be said about student achievement versus student learning and human development. The Finnish have done away with standardized testing and have based their education on the follow beliefs (Sahlberg, 2007):

1. Some basic school for all students
2. Well-trained teachers (almost all Finnish teachers have a master's degree)
3. Intelligent accountability
4. Culture of trust
5. Sustainable leadership

These principles have led the Finnish to do away with standardized testing and also students and teacher have a culture of learning to know, to do, and to encourage natural curiosity. Further, teachers have more freedom in curriculum planning, which allows them to spend more time providing students with basic knowledge of problem solving and reasoning. Stackett et al. (2008) found that while high-stakes testing provide consistent research, it does not necessarily portray accurate results of student learning. Therefore, while this study did not show a significant difference in student achievement, it did not take into consideration student learning or human development.

Some of the many reasons that ability grouping is implemented is to meet the demands of high-stakes testing, meet individual student needs, be able to teach appropriate skills for students at each level, and reducing the boredom of advanced students. Yet another reason is to make it easier for teachers to teach a group of students at a similar level of instruction (Forgasz, 2010;

Hornby et al., 2011). While the above reasons are targeting the needs of students, research found that more often than not grouping students into classrooms based on ability does not show a significant difference versus those who are in a mixed-ability classroom.

Recommendations for Practice

Based on the findings of this study, the author provides the following recommendations for practice:

1. When making administrative decisions that impact the instructional setting administrators should take into consideration both the academic and social impact of the instructional practice. It is imperative that the whole child be considered, rather than just his or her academic growth.
2. Teachers and administrators should meet to discuss the implications of grouping students by ability and how that would impact their students based on the individual school demographics. Each school is its own entity and must take into consideration outside factors that may impact student learning.
3. Administrators should consider the social impact that grouping students by ability will have prior to implementing the practice. The social and academic benefits as well as risks must be taken into consideration when making instructional decisions.
4. Considering the impact that teachers have on student achievement, teachers should be involved in the decision-making regarding ability grouping and how it would impact their teaching. If teachers are on-board with instructional strategies, they will have a larger impact on student learning. Teachers have the ability to have a large impact on student

learning and, therefore, must be on-board with instructional strategies that are implemented.

Implications for Further Research

The results of this study have prompted the researcher to make the following recommendations for further research regarding student achievement and the type of grouping implemented in the classroom:

1. Conduct similar studies with a larger number of schools and in different regions to determine if the results are similar to the results of this study.
2. Conduct research regarding individual student achievement in each of the four proficiency levels based on the type of grouping implementing to determine if there is a difference in individual students versus schools as a whole.
3. Investigate the difference between student growth scores and student achievement between students receiving instruction in classrooms grouped by ability and those not grouped by ability.
4. Use a mixed-methods approach to explore the social impact that ability grouping has on students and how student self-concept impacts achievement on standardized assessments.
5. Conduct research to determine the relationship between teacher effectiveness, student achievement, and ability grouping to assess if there is a difference between groupings based on teacher effectiveness.
6. Use qualitative research methods to assess teacher perception of ability grouping and how that impacts their teaching strategies.

7. Conduct studies regarding ability grouping and student achievement taking into consideration Title I status.
8. Investigate the difference in student growth and achievement between students participating in a specific intervention or gifted and talented program and students receiving intervention and acceleration within a self-contained classroom.
9. Conduct research in which student learning and human development are assessed rather than student achievement in both ability grouped settings and settings not grouped by ability.

Conclusion

School administrators will continue to make decisions regarding instructional practices based on research results, best practices, and personal beliefs. It is imperative that administrators understand the mixed results in research regarding grouping students by ability for reading and math instruction as well as the social impact (Ireson et al., 2001). The results of this study will add to the numerous other studies with similar results indicating that there is not a significant difference in the percentage of students scoring in each proficiency level in grades 3, 4, and 5 in both reading and math on the TCAP. This study collected data from public schools in 11 East Tennessee Counties. The schools were asked to self-report the type of grouping that was implemented in their schools for the 2012-2013 school year: ability grouped classrooms or classrooms not grouped by ability. Overall, there was not a statistically significant difference in grade 3 in reading or math in any of the proficiency levels (below basic, basic, proficient, and advanced) between schools that implemented ability grouped classrooms and those that did not

group by ability. Further, there was no significant difference in grades 4 or 5 in reading and math in any of the proficiency levels between the two types of classroom settings. When contacted many schools that implemented classroom settings with no ability grouping for reading and math still did participate in grouping within the classroom in order to differentiate instruction. As a whole, the results of this study are consistent with many others in finding that there is no significant difference achievement scores based on placing students in an ability grouped setting or in a classroom that does not group by ability (Reuman, 1989; Wouters et al., 2012).

REFERENCES

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: National Council on Measurement in Education.

Armstrong, T. (2006). *The best schools: How human development research should inform educational practice*. http://www.ascd.org/publications/books/106044.aspx.

ASCD. (n.d.). *A lexicon of learning*. Retrieved June 11, 2014, from ASCD: http://www.ascd.org/Publications/Lexicon-of-Learning/A.aspx

Aslanargun, E. (2012). Principals' values in school administration. *Educational Sciences: Theory and Practice, 12*(2), 1339-1344.

Ansalone, G. (2010). Tracking: educational differentiation or defective strategy. *Educational Research Quarterly, 34*(2), 3-17. Retrieved from https://login.ezproxy.etsu.edu:3443/login?url=http://search.proquest.com/docview/14170 50908?accountid=10771

Betts, J. R. (2011) The Economics of tracking in education. *Handbook of the Economics of Education.* 3. 341-381.

Boaler, J. (2008). Promoting 'relational equity'and high mathematics achievement through an innovative mixed-ability approach. *British Educational Research Journal*, *34*(2), 167-194.

Braddock, J. H., & Slavin, R. E. (1992). *Why ability grouping must end: Achieving excellence and equity in American education*. [PDF] John Hopkins University Center for Research on Effective Schooling for Disadvantages Students.

Brulles, D., Saunders, R., Cohn, S. (2010). Improving performance for gifted students in a cluster grouping model. *Journal for the Education of the Gifted.* 34(2). 327–350.

Bui, S., Imberman, S., & Craig, S. (2012). Poor results for high achievers. *Education Next, 12*(1) Retrieved from https://login.ezproxy.etsu.edu:3443/login?url=http://search.proquest.com/docview/12378 26558?accountid=10771

Calik, T., Sezgin, F., Kavgaci, H., & Cagatay Kilinc, A. (2012). Examination of relationships between instructional leadership of school principals and self-efficacy of teachers and collective teacher efficacy. *Educational Sciences: Theory and Practice*, 12(4), 2498-2504

Catsambis, S., & Buttaro, A. (2012). Revisiting "kindergarten as academic boot camp": A nationwide study of ability grouping and psycho-social development. *Social Psychology of Education: An International Journal, 15*(4), 483-515. Retrieved from

https://login.ezproxy.etsu.edu:3443/login?url=http://search.proquest.com/docview/1312424223?accountid=10771

Catsambis, S., Mulkey, L., Buttaro, A., Steelman, L., & Koch, P. (2012). Examining gender differences in ability group placement at the onset of schooling: The role of skills, behaviors, and teacher evaluations. *Journal Of Educational Research*, *105*(1), 8-20.

Cheng, W. Y., Lam, S. F., & Chan, C. Y. (2008). When high achievers and low achievers work in the same group: The roles of group heterogeneity and processes in project-based learning. *British Journal of Educational Psychology*, *78*(2), 205-221.

Chmielewski, A. K., Dumont, H., & Trautwein, U. (2013). Tracking effects depend on tracking type: An international comparison of students' mathematics self-concept. *American Educational Research Journal*, *50*(5), 925-957.

Chorzempa, B., & Graham, S. (2006). Primary-grade teachers' use of within-class ability grouping in reading. *Journal Of Educational Psych* (ASCD)*ology*, 98(3), 529-541.

Christie, K. (2008). Dat's story: things have got to change. *Phi Delta Kappan*, *89*(7), 469-544.

Duflo, E., Dupas, P., & Kremer, M. (2009). Can tracking improve learning? *Education Next , 9* (3), 64-70.

Cotle,T.L.(1974).What tracking did to Olie Taylor. *Social Policy*, 5,21-24.

Duflo, E.; Dupas, P.; & Kremer, M. (2008). Peer effects and the impact of tracking: evidence from a randomized evaluation in Kenya. *California Center for Population Research*. UCLA: California Center for Population Research. Retrieved from: https://escholarship.org/uc/item/9607k4rf

Elmore, R. F. 2003. A plea for strong practice. *Education Leadership* 61 (3): 6-10.

Emery, K. (2007). Corporate control of public school goals: High-stakes testing in its historical perspective. *Teacher Education Quarterly, 34*(2), 25-44. Retrieved from https://login.ezproxy.etsu.edu:3443/login?url=http://search.proquest.com/docview/61947501?accountid=10771

Fendler, L., & Muzaffar, I. (2008). The history of the bell curve: Sorting and the idea of normal. *Educational Theory, 58*(1), 63-82. Retrieved from https://login.ezproxy.etsu.edu:3443/login?url=http://search.proquest.com/docview/214138034?accountid=10771

Flashman, J. (2012). Academic achievement and its impact on friend dynamics. *Sociology of Education, 85*(1), 61-80. Retrieved from https://login.ezproxy.etsu.edu:3443/login?url=http://search.proquest.com/docview/918717204?accountid=10771

Fram, M., Miller-Cribbs, J. E., & Van Horn, L. (2007). Poverty, race, and the contexts of achievement: Examining educational experiences of children in the U.S. South. *Social Work*, 52(4), 309-319.

Forgasz, H. (2010). Streaming for mathematics in victorian secondary schools. *Australian Mathematics Teacher, 66*(1), 31-40. Retrieved from https://login.ezproxy.etsu.edu:3443/login?url=http://search.proquest.com/docview/742865225?accountid=10771

Gallagher, S., Smith, S., & Merrotsy, P. (2011). Teachers' perceptions of the socioemotional development of intellectually gifted primary aged students and their attitudes towards ability grouping and acceleration. *Gifted & Talented International*, *26*(1/2), 11-24.

Green, S. B., & Salkind, N. J. (2011). *Using SPSS for Windows and Macintosh: Analyzing and understanding data.* Upper Saddle River, NJ: Prentice Hall.

Good, T. L. (1997). Do students learn more in heterogeneous or homogeneous groups?. *DOCUMENT RESUME ED 268 075 SP 027 044*, 27.

Guisbond, L., & Neill, M. (2004). Failing our children: No Child Left Behind undermines quality and equity in education. *Clearing House*, *78*(1), 12-16.

Hallam, S., & Ireson, J. (2003). Secondary school teachers' attitudes towards and beliefs about ability grouping. *British Journal Of Educational Psychology*, 73(3), 343-356.

Hallam, S., & Ireson, J. (2008). Subject domain differences in secondary school teachers' attitudes towards grouping pupils by ability. *Zbornik Instituta za pedagoška istraživanja*, *40*(2), 369-387.

Hallam, S., & Parsons, S. (2013). Prevalence of streaming in UK primary schools: Evidence from the millennium cohort study. *British Educational Research Journal, 39*(3), 514-544. Retrieved from https://login.ezproxy.etsu.edu:3443/login?url=http://search.proquest.com/docview/1509088028?accountid=10771

Hong, G., Corter, C., Hong, Y., & Pelletier, J. (2012). Differential effects of literacy instruction time and homogeneous ability grouping in kindergarten classrooms: Who will benefit? who will suffer?. *Educational Evaluation & Policy Analysis*, *34*(1), 69-88.

Hornby, G., Witte, C., & Mitchell, D. (2011), Policies and practices of ability grouping in New Zealand intermediate schools. *Support for Learning*, 26: 92–96.

Houchens, G.W., & Keedy, J.L. (2009). Theories of practice: Understanding the practice of educational leadership. *Journal of Thought*. 44(3-4), 49-61.

Hung, W., Smith, T. J., Harris, M. S., & Lockard, J. (2010). Development research of a teachers' educational performance support system: The practices of design, development, and evaluation. *Educational Technology Research & Development*, 58(1), 61-80.

Ireson, J., & Hallam, S. (1999). Raising standards: Is ability grouping the answer?. *Oxford Review of Education*, *25*(3), 343-358.

Ireson, J., Hallam, S., & Plewis, l. (2001). Ability grouping in secondary schools: Effects on pupils' self-concepts. *British Journal Of Educational Psychology*, 71(2), 315.

Ireson, J., & Hallam, S. (2005). Pupils' liking for school: Ability grouping, self-concept and perceptions of teaching. *British Journal Of Educational Psychology*, 75(2), 297-311.

Ireson, J., & Hallam, S. (2009). Academic self-concepts in adolescence: Relations with achievement and ability grouping in schools. *Learning and Instruction*, *19*(3), 201-213.

Kemp, S., & Watkins, D. (1996). Self-esteem and academic streaming in Hong Kong. *Journal Of Social Psychology*, 136(5), 651-653.

Kintz, M. (2011). Ability grouping and how it is affecting American classrooms. *ESSAI*, *9*(1), 20.

Kulik, J. A. (1992). *An analysis of the research on ability grouping: Historical and contemporary perspectives. research-based decision making series* National Research Center on the Gifted and Talented, The University of Connecticut, 362 Fairfield Rd, U-7, Storrs, CT

Kulik, J. A., & Kulik, C.-L. C. (1987). Effects of ability grouping on student achievement. *Equity and Excellence*, 23, 22–30.

Kususanto, P., Ismail, H. N., & Jamil, H. (2010). Students' self-esteem and their perception of teacher behavior: A study of between-class ability grouping. *Electronic Journal of Research in Educational Psychology, 8*(2), 707-724.

Lee, J. (2008). Is test-driven external accountability effective? Synthesizing the evidence from cross-state causal-comparative and correlational studies. *Review of Educational Research, 78*(3), 608-644. Retrieved from https://login.ezproxy.etsu.edu:3443/login?url=http://search.proquest.com/docview/214113999?accountid=10771

Lleras, C., & Rangel, C. (2009). Ability grouping practices in elementary school and African American/Hispanic achievement. *American Journal of Education*, *115*(2), 279-304.

Lou, Y. (1996). *Within-class grouping: A meta-analysis*. Retrieved from https://login.ezproxy.etsu.edu:3443/login?url=http://search.proquest.com/docview/62603521?accountid=10771

McMillan, J., & Schumacher, S. (2010). *Research in education: Evidence-based inquiry.* Boston, MA: Pearson.

Mokhtar, I. A., Majid, S., & Foo, S. (2008). Teaching information literacy through learning styles: The application of Gardner's multiple intelligences. *Journal of Librarianship and Information Science*, *40*(2), 93-109.

Musoleno, R. R., & White, G. P. (2010). Influences of high-stakes testing on middle school mission and practice. *RMLE Online: Research in Middle Level Education*, *34*(3), 1-10.

Neihart, M. (2007). The Socioaffective impact of acceleration and ability grouping recommendations for best practice. *Gifted Child Quarterly*, *51*(4), 330-341.

Oakes, J. (2008). Keeping track: Structuring equality and inequality in an era of accountability. *The Teachers College Record*, *110*(3), 700-712.

Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Pierce, R. L., Cassady, J. C., Adams, C. M., Speirs Neumeister, K. L., Dixon, F. A., & Cross, T. L. (2011). The effects of clustering and curriculum on the development of gifted learners' math achievement. *Journal For The Education Of The Gifted*, *34*(4), 569-594.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational psychologist*, *37*(2), 91-105.

Plunkett, M., & Kronborg, L. (2011). Learning to be a teacher of the gifted: The importance of examining opinions and challenging misconceptions. *Gifted & Talented International*, *26*(1/2), 31-46.

Popham, W. (2011). Assessment literacy overlooked: A teacher educator's confession. *Teacher Educator*, *46*(4), 265-273.

Popham, W. (2007). Instructional insensitivity of tests: Accountability's dire drawback. *Phi Delta Kappan*, *89*(2), 146-155.

Puzio, K., & Colby, G. (2010). *The effects of within class grouping on reading achievement: A meta-analytic synthesis.* Society for Research on Educational Effectiveness. 2040 Sheridan Road, Evanston, IL 60208.

Reuman, D. A. (1989). How social comparison mediates the relation between ability-grouping practices and students' achievement expectancies in mathematics. *Journal Of Educational Psychology*, 81(2), 178-189.

Riffert, F. (2005). The use and misuse of standardized testing: A Whiteheadian point of view. *Interchange*, *36*(1/2), 231-252.

Robinson, J. (2008). Evidence of a differential effect of ability grouping on the reading achievement growth of language-minority Hispanics. *Educational Evaluation and Policy Analysis June 2008, Vol. 30, No. 2, pp. 141–180*

Sackett, P. R., Borneman, M. J., & Connelly, B. S. (2008). High stakes testing in higher education and employment: Appraising the evidence for validity and fairness. *American Psychologist*, *63*(4), 215.

Sahlberg, P. (2007). Education policies for raising student learning: The Finnish approach. *Journal of Education Policy*, *22*(2), 147-171.

Schofield, J. (2010). International evidence on ability grouping with curriculum differentiation and the achievement gap in secondary schools. *The Teachers College Record*, *112*(5), 8-9.

Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research* , 293-336.

Slavin, R. E. (1990). Achievement effects of ability grouping in secondary schools: A best-evidence synthesis. *Review of educational research*, *60*(3), 471-499.

Solomon, Y. (2007). Experiencing mathematics classes: Ability grouping, gender and the selective development of participative identities. *International Journal of Educational Research*, *46*(1), 8-19.

Stewart, W. (2013). Do setting and streaming work? *The Times Educational Supplement Scotland,* (2311), 24. Retrieved from https://login.ezproxy.etsu.edu:3443/login?url=http://search.proquest.com/docview/1356891434?accountid=1077

Tennessee Department of Education. *Assessment.* Retrieved June 26, 2014, from TDOE: http://www.tn.gov/education/assessment/grades_3-8.shtml

Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O., & Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal Of Educational Psychology*, 98(4), 788-806.

Witte, R. S., & Witte, J. S. (2007). *Statistics.* Danvers, MA: John Wiley & Sons.

Wouters, S., De Fraine, B., Colpin, H., Van Damme, J., & Verschueren, K. (2012). The effect of track changes on the development of academic self-concept in high school: A dynamic test of the big-fish--little-pond effect. *Journal Of Educational Psychology*, 104(3), 793-805.

Zirkel, P. A. (1998). Testing the limits or limiting the tests?. *Phi Delta Kappan*, *80*(4), 330-331.

VITA

ANNA KATHRYN MCCARTER

Education:

Public Schools, Stevensville, Michigan

B.A. Communications, Carson-Newman University, Jefferson City, Tennessee, 2008

M.A. Teaching, Carson-Newman University, Jefferson City, Tennessee, 2009

Ed.D, East Tennessee State University, Johnson City, Tennessee, 2014

Professional Experience:

Teacher, Hamblen County Public Schools; Morristown, Tennessee, 2010-2014

Teacher, Sevier County Public Schools; Sevierville, Tennessee, 2014- present