

HOMOGENEOUS GROUPING AND ITS EFFECTIVENESS IN THE ELEMENTARY
SCHOOL SETTING

A Dissertation

Presented to

The Faculty of the Education Department

Carson-Newman University

In Partial Fulfillment

of the

Requirements for the degree

Doctor of Education

Angela Johnson

May 2016

Copyright © 2016 by Angela Johnson
All Rights Reserved.



Dissertation Approval

Student Name: Angela Johnson

Dissertation Title: HOMOGENEOUS GROUPING AND ITS EFFECTIVENESS IN THE ELEMENTARY SCHOOL SETTING

This dissertation has been approved and accepted by the faculty of the Education Department, Carson-Newman University, in partial fulfillment of the requirements for the degree, Doctor of Education.

Dissertation Committee:

Signatures: (Print and Sign)

A handwritten signature in cursive script that reads "Deborah Hayes".

Dr. Deborah Hayes
Dissertation Chair

A handwritten signature in cursive script that reads "Brenda Dean".

Dr. Brenda Dean
Methodologist Member

A handwritten signature in cursive script that reads "Mark N. Brock".

Dr. Mark Brock
Content Member

Approved by the Dissertation Committee

Date: April 8, 2016

I hereby grant permission to the Education Department, Carson-Newman University, to reproduce this research in part or in full for professional purposes, with the understanding that in no case will it be for financial profit to any person or institution.

Signature: Angela Johnson
Date: May 2016

Abstract

Homogenous Grouping and its Effectiveness in the Elementary School Setting

Angela Johnson

Department of Education, Carson-Newman University

May 2016

Homogeneous grouping is an educational method utilized to differentiate instruction as a way for students to obtain academic achievement. The objective for implementing homogeneous placement is to increase students' achievement by providing instruction at their present academic level. Despite the fact it is utilized often in education, there has been conflicting research in respect to the effect of homogeneous placement and its association with students' academic achievement. The primary goal of this research study was to ascertain if there was a statistically significant difference between elementary student academic achievement in grades three, four, and five based on the kind of placement in classroom settings. Data were gathered from the Tennessee Department of Education website and the Pearson Published Reports for the 2014-2015 school year, as well as from school principals concerning whether or not students were homogeneously placed. Once the data were gathered, a two-sample z-test was utilized to determine if there was a significant difference between students who received instruction in a homogeneous setting and those who did not. The results of this research study indicated that there was a significant difference in proportions of students in two of the four categories, in respect to the type of grouping implemented. A significant difference was found in the area of mathematics. Therefore, two of the four null hypotheses were rejected.

Acknowledgements

At this time, I would like to acknowledge all of my family, friends, and esteemed colleagues. I will always be thankful that God has allowed me to be associated with such wonderful people. I feel blessed for your support and encouragement, as I have continued to travel this journey.

I will always be grateful to the Department of Educational Leadership and Policy Analysis, faculty, and staff of Carson Newman University for the opportunity to learn under such wise leadership. I appreciate you for helping me continually grow as an educator. The encouragement and guidance I have received through this program have been pivotal in my career.

I am especially grateful to my distinguished committee members, Dr. Deborah Hayes, Dr. Brenda Dean, and Dr. Mark Brock. I sincerely appreciate your support, encouragement, and guidance as I have progressed through the dissertation process. I would like to thank Dr. Julia Price and Dr. Ernest Walker for all of their help and support over the past three years as well. I would also like to thank my classmate, Christy Walker, for providing me a great deal of help and support. Moreover, special thanks to my family, especially my husband and aunt, Kathy Hodge, for their good-natured attitude in respect to this process and for their pride in this accomplishment. Without a doubt, it was a team effort.

Dedication

To my husband and best friend, Kerry, you have always encouraged me in all endeavors. I am thankful that God allowed us to be together. I am so thankful for your constant love and support. I would also like to recognize my grandparents, George and Dorothy Johnson, who always provided me with love and support. I will always be grateful to them for giving me a loving Christian home. I will forever be thankful for the example that my grandparents have provided. Even though my grandfather is no longer with us, I know that he is with us in spirit. Last, but certainly not least, I would like to thank my Lord and Savior, Jesus Christ.

Table of Contents

Abstract.....	v
Acknowledgements.....	vi
Dedication.....	vii
List of Figures.....	x
Chapter One: Introduction	1
Statement of Problem.....	2
Significance of Study.....	3
Research Questions.....	4
Limitations and Delimitations.....	5
Definition of Terms.....	5
Organization of Document.....	8
Chapter Two: Review of Literature	10
Historical Context.....	13
Positive Academic Effect of Homogeneous Grouping	15
Negative Academic Effect of Homogeneous Grouping	16
Homogeneous Grouping in Regard to High Achievers	18
Homogeneous Grouping in Mathematics and ELA.....	19
Academic Effect in Mathematics.....	21
Academic Effect in English Language Arts.....	23
Effect of Homogeneous Grouping on Student Self-esteem & Inequality.....	25
Effect of Homogeneous Grouping on Student Self-esteem.....	26
Effect of Homogeneous Grouping on Social Inequality.....	28
Standardized Assessment and Learning.....	30
Measuring Student Learning.....	32
Limitations of Standardized Assessment.....	35
Summary.....	37
Chapter Three: Methods and Procedures	38
Research Questions.....	38
Instrumentation	39
Research Participants and Setting for the Study.....	40
Population.....	40
Setting	40
Data Collection Procedures.....	40
Validity, Reliability, Bias	41
Human Safety.....	42
Ethical Issues.. ..	42
Data Analysis Procedures	42
Summary.....	43

Chapter Four: Analysis of Data	44
Analysis of Research Questions.....	44
Summary.....	52
Chapter Five: Conclusion, Implication, and Recommendations	54
Summary of Results.....	54
Implications for Practice.....	56
Recommendations for Practice.....	58
Implications for Further Research.....	59
Conclusion.....	60
References	62
Appendices	72
Appendix A: Request to Conduct Research.....	73
Appendix B: Approval Response.....	75
Appendix C. Initial Email to Participants.....	77
Appendix D. Questions for Participants.....	79

List of Figures

Figure 4.1: Percentage of Students Scoring in the Proficient Level in Mathematics	47
Figure 4.2: Percentage of Students Scoring in the Below Proficient Level in Mathematics...	49
Figure 4.3: Percentage of Students Scoring in the Proficient Level in ELA	51
Figure 4.4: Percentage of Students Scoring in the Below Proficient Level in ELA.....	53

CHAPTER ONE

Introduction

Homogeneous grouping has been defined as “placing students in classrooms based on their current academic ability level in a certain subject” (Davidson, 2009). In the United States, students have been placed into ability groups based on their academic strengths and weaknesses (Davidson, 2009; Slavin, 1987). In some educational settings, students were placed into classrooms that were not grouped homogeneously, and educators gave group instruction in respect to the specific individual need of students. In some instructional settings, students were placed into homogeneously-grouped classes based on academic ability or academic achievement (Emery, 2007). Though homogeneous grouping is a common educational practice, there has been a great deal of research that supports both classrooms that are heterogeneously grouped and those that are homogeneously grouped (Campbell, 2014; Davies, Hallam & Ireson, 2003). Regardless of the educational practice implemented, the objective of placing students into instructional groups in respect to their capability level is to increase the academic success of students (McCarter, 2014).

With the focus placed on the importance of instructional data, teachers across the country are utilizing changes in their classrooms to face the growing demands to improve student academic achievement, as measured by standardized test scores (Sanders, Wright, & Horn, 1997). In 2001, President George W. Bush signed the No Child Left Behind (NCLB) Act into legislation. A section of the act emphasized bridging the academic achievement gap in schools

across the United States based on its current academic achievement scores. The No Child Left Behind Act, the reauthorization of the Elementary and Secondary Education Act, supports the idea of implementing rigorous standards and establishing objectives that can be measured (Klein, 2015).

In education, homogeneous grouping has produced positive and negative effects on students academically, socially, and emotionally (Reuman, 1989). Over the past three decades, the utilization of homogeneous grouping in the United States has been a very controversial topic (Hornby, Witte, & Mitchell, 2011). While Duflo, Dupas, and Kremer (2009) indicated academic benefits to placing students in homogeneous classrooms, other research indicated potential for decreased self-efficacy in those who experienced this type of academic placement (Reuman, 1989). With this in mind, teachers are responsible not only for providing a vigorous academic experience for their students, but also for establishing successful members of society. For this reason, teachers must extensively evaluate the influence of their instructional choices.

Statement of Problem

Due to the changes in legislation in the United States, teachers have utilized various instructional methods to face the demands placed upon them (Hornby et al., 2011). Historically, academic capabilities and chosen vocations categorize students for homogeneous grouping and placement into instructional tracks. However, changes in legislation have shifted the motive for homogeneous placement from classes that will prepare students for future vocations to utilizing standardized data for placement in classes in which a student will flourish academically (Emery, 2007). With conflicting research studies in respect to the academic benefits of grouping students based on academic capability (Duflo et al., 2009), educators have also evaluated the social impact of homogeneous grouping. After further analysis, some researchers have indicated that

there are negative effects associated with homogeneous placement and student self-efficacy (Barker Lunn, 1970; Catsambis & Buttaro, 2012; Kintz, 2011).

In education, placing students into classrooms based on academic achievement is a common response to providing differentiated instruction for students. At the beginning of one's school career, students in the United Kingdom are placed into tracks based on standardized assessment scores (Ireson & Hallam, 2001). In many Asian countries, students are placed into specific schools based on their performance on standardized assessments as well. In the United States, there are various methods in respect to ability grouping (Catsambis & Buttaro, 2012; Fendler & Muzaffar, 2008; Kintz, 2011). Upon entering high school, students often follow either a vocational or a university track. However, at this time, limited extensive research has been conducted concerning the influence homogeneous placement has on academic achievement at the elementary school level (Matthews, Ritchotte, McBee, 2013).

This study was conducted as an evaluation of student academic achievement on the Tennessee Comprehensive Assessment Program (TCAP) in mathematics and English language arts for grades three, four, and five for the 2014-2015 school year. It compared students who received instruction in homogeneously-grouped classrooms to those who received instruction in a classroom not homogeneously-grouped. The objective of this research study was to ascertain if there was a statistically significant difference among student academic achievement in grades three, four, and five based on the type of instructional placement.

Significance of the Study

Across the United States, school administrators and teachers are faced with the daunting task of making important decisions concerning the instructional methods that influence learning. In mathematics and English language arts, one of the most frequently used methods is placing

students into classrooms based on their academic capability. To make informed decisions that indicate research-based methods, it was important to ascertain if there was a significant difference in the scores of students in the Tennessee Comprehensive Assessment Program (TCAP), Proficient or Below Proficient categories, within schools that placed students into classrooms for instruction based on their capability, and those that placed students into classrooms with students of varying capability. TCAP data produce four proficiency levels and for the purpose of this study advanced and proficient are combined in the category of proficient and basic and below basic are combined in the category entitled proficient. This study has provided an evaluation of the difference in Tennessee Comprehensive Assessment Program (TCAP) assessment scores according to the type of placement implemented in schools. The objective of this study was to provide educators with additional research to inform decision making.

Research Questions

This study was guided by the following research questions:

Research Question 1

Is there a significant difference in the scores of students in the Tennessee Comprehensive Assessment Program (TCAP) levels: Proficient and Below Proficient for grades three, four, and five in mathematics between schools that group students homogeneously and schools that do not?

Research Question 2

Is there a significant difference in the scores of students in the Tennessee Comprehensive Assessment Program (TCAP) levels: Proficient and Below Proficient for grades three, four, and

five in English language arts between schools that group students homogeneously and schools that do not?

Limitations/Delimitations

- The research data focused on elementary school criterion-referenced assessment achievement scores found on the Tennessee Department of Education website and the school district office.
- The classification of grade level placement was limited to the administrator's answers when contacted in respect to homogeneous grouping at each school.
- The elementary schools were categorized based on emails placed to school administrators, and schools were coded to ensure anonymity.
- The research data were restricted to the 2014-2015 school year.
- The accuracy of coding the type of homogeneous grouping was based on the administrators' understanding of grouping definitions and accurately reporting the information. The principal's perception regarding grouping could impact the type of classroom design at the respective schools.
- The research data were limited to eight Title I schools located in a rural county in East Tennessee.
- The research data were limited to scores based on the Tennessee Comprehensive Assessment Program.

Definition of Terms

Ability Grouping. Within this study, the definition of ability grouping was, based on achievement scores, students are placed into various classes in respect to their current academic

ability level in a specific subject (also known as homogeneous grouping) (Marzano et. al., 2008, McCarter, 2014).

Achievement Gap. For this study, the definition of achievement gap was, the lasting variation in academic attainment among subgroups of students as specified by standardized test scores, class grades, and other academic data (Marzano et al., 2008; McCarter, 2014).

Achievement Growth. For this study, the definition of achievement growth was, the academic progress made over a certain period of time. It can be tracked for the individual students, schools, states, or countries, and a wide variety of variances may be used to determine whether academic growth has been accomplished (Sanders, Wright, & Horn, 1997).

Achievement Tests. For this study, the definition of achievement tests was the methods that are applied to anticipate student learning in various subjects. These standardized, multiple-choice assessments are utilized to determine students' academic achievement in subjects found in most school districts' curriculum. The test results are used to compare students and schools with others that are in the same county, across the state, and across the country” (Marzano et al., 2008; McCarter, 2014, p. 35).

Banding. For this study, the definition of banding was the educational practice of placing students into classes or groups in respect to one’s academic capability (Wouters, DeFraine, Colpin, Van Damme, & Verschuren, 2012; McCarter, 2014).

Clustering. For this study, the definition of clustering was the educational practice of placing students of a particular academic capability in a heterogeneous classroom and providing differentiation in instruction (Pierce, Cassady, & Adams, 2011).

Criterion-referenced. Within this study, the definition of criterion-referenced test is based on specific predetermined educational outcomes. Students' scores are not compared to one another but are based on individual performance.

Heterogeneous Grouping. Within this study, the definition of heterogeneous grouping was educational practice of placing students of various capabilities in the same classroom for academic instruction (Marzano et al., 2008; McCarter, 2014).

Homogeneous Grouping. Within this study, the definition of homogeneous grouping was the educational practice of placing students in classes based on their academic capability (Marzano et al., 2008; McCarter, 2014).

Gifted. For the purpose of this study, the definition of gifted or talented was students who give evidence of high academic achievement capability in such areas as intellectual, creative, artistic, or leadership skills (Colangelo & Davis, 1997).

Global Community. For the purpose of this study, the definition of global community was the people or nations of the world, considered as being closely connected by modern telecommunications and as being economically, socially, and politically interdependent (Willard, 1996).

Regrouping. For the purpose of this study, the definition of regrouping was, a system whereby students are separated into various groups or classes according to test scores or relative scholastic capability (Wouters et al., 2012).

Standardized Intelligence Tests. Within this study, the definition of standardized intelligence tests was psychological assessments designed to measure a wide variety of mental functions, such as reasoning, comprehension, and judgment. The main objective of the intelligence tests is to ascertain an idea of one's intellectual potential (Sattler, 2001).

Setting. For the purpose of this study, the definition of setting was educational practice of placing students between classes based on academic capability for each branch of study. This is the most adjustable form of within subject grouping, and it is possible for a student to be in various groups for all subjects (Wouters et al., 2012; McCarter, 2014).

Streaming. For the purpose of this study, the definition of streaming was the educational practice of placing students within subjects according to one's capability (Wouters et al., 2012; McCarter, 2014).

Tracking. For the purpose of this study, the definition of tracking was the method of separating students for academic instruction based on their capabilities. Students are placed on a particular track such as college-bound, general, remedial and vocational, and are given a curriculum that varies according to their perceived abilities and future positions in life. Within the elementary school setting, the method is called grouping (Marzano et al., 2008; McCarter, 2014).

Organization of the Document

This research study was separated into five chapters. Chapter One provided an introduction to the study. The first chapter contained the context and background, outline of the problem, and identified the research question. In addition, the theoretical and conceptual frameworks for the study were identified, terms were defined, and limitations for the study were outlined. The significance of the study was also discussed. Chapter Two contained a review of literature. The review of literature was an analysis related to the topics of homogeneous grouping and its effects on the academic achievement of students. The research methodology was found in Chapter Three. The third chapter included descriptions of the population, instruments, and research procedures used to conduct the study. A description of data analysis procedures was also included. Chapter Four contained the analysis of data in respect to this research study. In

Chapter Five, the summary of results, implications and recommendations for practice, implications for further research, and conclusions were stated.

CHAPTER TWO

Review of Literature

The objective of this research study was to ascertain if there was a statistically significant difference between student academic achievement in grades three, four, and five, according to the type of placement in instructional settings. This review of literature discussed homogeneous grouping, the impact of homogeneous grouping on student's self-esteem, and the effects of homogeneous grouping on student's academic performance. In order to provide a framework for the research study, specific emphasis was placed on the overall effect of homogeneous placement.

Homogeneous grouping, also known as ability grouping, is the educational method of placing students into groups in respect to their academic achievement level (Slavin, 1990). In the United States, educators have used the method of homogeneous grouping in secondary schools by placing students into vocational or university tracks based on their chosen pursuit and academic capabilities. As standardized assessments have become more instrumental in the educational setting, teachers have applied homogeneous grouping as a way to increase the academic achievement of students. Despite its common utilization, researchers have found contrasting evidence in respect to the effectiveness of homogeneous grouping. Furthermore, Hornby, Witte, & Mitchell (2011) indicated that homogeneous placement had a negative impact on the academic achievement for students in English language arts and had little effect on students in mathematics in respect to standardized test scores.

To meet the academic needs of students and increase academic growth, educators have utilized homogeneous placement as a common practice. As indicated by Chorzempa and Graham (2006), homogeneous grouping was the educational method of placing students into groups in regard to their academic achievement level. However, research studies have indicated the kind of homogeneous grouping makes a difference in student development all through the school year (Reuman, 1989). Even though homogeneous grouping has been a common educational practice in the United States, researchers have indicated that homogeneous grouping is a societal injustice and lowers the self-efficacy of students as well as their academic achievement (Fram, Miller-Cribbs, & Van Horn, 2007).

Educators have increasingly maintained and applied stratification practices such as homogeneous grouping, streaming, and setting in order to raise levels of academic achievement (Trigg-Smith, 2011). According to the research study of Trigg-Smith (2011), the researcher noted how educational policies have contributed to the decisions of educators in respect to ability grouping. The researcher opined existing theories of intelligence and ability grouping have reinforced the methodology, the plausible frameworks for the exploration of equality in ability grouping, and the importance of the impact that teachers have on students.

Research has been conducted to determine if social inequality occurs when the practice of homogeneous grouping is implemented. According to existing research, data has verified that homogeneous grouping benefits advanced students more than other groups (Hornby, Witte, & Mitchell, 2011). In agreement with Hornby, Witte and Mitchell (2011), Preckel and Brull (2008) indicated that homogenous grouping benefits advanced students more than other groups; however, the researchers pointed out that homogeneous grouping should be conducted full-time for there to be a significant difference in academic gains. Based on a study by Catsambis,

Mulkey, Buttaro, and Steelman (2011), the researchers found that boys and girls were equally placed into average groups in kindergarten; however, boys were not represented as often in higher achieving groups, and girls were not represented as often in lower achieving groups.

Kulik's (1992) meta-analysis focused on the amount of growth found in homogeneous grouping. Kulik indicated the following affected the amount of student academic achievement that occurred: classes should be conformed to the same instructional guidelines. The researcher noted that the guidelines should be conformed to the instructional curriculum in respect to the academic capability of students and should be conformed to the curriculum for the special needs and academically talented students. Therefore, when classifying students, Kulik (1992) concluded that the amount of adjustments made to the instructional curriculum has affected the academic attainment. The highly debated topic of homogeneous grouping has brought forth a substantial amount of research that both supports and discredits homogeneously-grouped classes.

The technique for homogeneous grouping has been frequently promoted in training as an approach to differentiated instruction. In spite of its prominent use, there are educational practices that take into account society and shifting vocabulary that encompass this educational topic. In the United Kingdom, ability grouping has been referred to as streaming or tracking, which is a form of between-class homogeneous grouping based on general academic ability. Banding is similar to streaming; however, it is more flexible and students are regrouped based on the particular subject area. Regrouping, also known as setting as an educational practice, is the least restrictive, and students are grouped based on their academic ability in each respective content area (Ireson & Hallam, 1999, Wouters et al., 2012).

Historical Context

Homogeneous grouping and its various adaptations are among the predominant educational practices of public schools in the United States. From a historical perspective, homogeneous grouping has been a common practice in the United States for over a hundred years (Rubin, 2008). During the Progressive period from 1890-1920, placing students into groups in public schools became a popular practice in educational systems throughout the country. During this time, local school boards often persuaded schools to place students into either a management track or vocational track (Emery, 2007). With the growth of immigrants to the United States in need of learning English, educators were prompted to place students into ability groups according to their chosen career paths.

In order to meet the needs of the industrialization era, schools across the United States were intent on classifying students in a more ordered fashion (Emery, 2007; McCarter, 2014). The standardized intelligence test was created to effectively classify students. Throughout the history of the American educational system, there have been changes toward a national curriculum and assessment. Homogeneous grouping became popular in the United States when ability grouped classrooms were often practiced in the 1990's. In recent years, placing students according to their academic ability has been utilized as a response to the increase in school accountability and to meet the competitive global community (Ireson & Hallam, 1999).

Despite its early popularity, the educational practice of tracking students ended in the United States in the late 1960's. Currently in European and Asian school systems, students are assigned to university or vocational tracks that indicate coursework through high school (Loveless, 1998). In many countries in Europe and Asia, based on assessment scores, students are selected into educational institutions with various post-secondary destinations (Smithers &

Robinson, 2013). In the early part of the 19th century, examination-based selections into high schools were a common educational practice in the United States. However, the comprehensive high school, with all students of a community attending the same school and then divided into tracks, has been the most commonly practiced method in education today.

The educational practices of homogeneous grouping and tracking attempted to match students with instruction focused on students' current academic achievement (Loveless, 2013). Despite the similarities, the methods vary in many ways. Often implemented on the high school level, tracked students were assigned to various classrooms and received instruction by different educators. However, in the United States, some high schools do not implement tracking. Loveless (2013) noted that some high schools across the country have implemented heterogeneous classrooms instead hoping to obtain academic success.

In the educational setting, there have been various motivating factors for categorizing students according to their academic capability. School principals have often encouraged their teachers to customize instruction as a way to improve the academic achievement of students (Hornby, Witte, & Mitchell, 2011). As noted by Forgasz (2010), another motivating factor for applying homogeneous grouping has been to help educators instruct a class of students at the same academic level.

Lou, Spence, Poulsen, Chambers, and Apollonia (1996) reported that homogeneous placement has been beneficial in respect to increasing the academic achievement of students. However, the researchers indicated that, for homogeneous placement to have a positive impact, the integration of instructional resources has been of the utmost importance. As noted by Lou, et al. (1996), the practice of homogeneous grouping has been commonly utilized in the field of

education as a method to differentiate instruction. Despite its popular use, there are variations based on culture and vocabulary that are related to this educational topic.

Positive Academic Effect of Homogeneous Grouping

When considering the positive effects of homogeneous grouping in education, Chorzempa & Graham (2006) noted the most widely supported purpose for implementing this form of placement was to provide academic instruction for all students that was affective and met the individual need of the student. A study conducted by Chorzempa and Graham (2006) involving 225 educators revealed 66% reported utilizing homogeneous placement. In addition, 69% of that 66% utilized this practice to satisfy the instructional needs of students. Thus, the findings of this study indicated that elementary educators believed homogeneous grouping aided students by meeting instructional needs.

Research studies have indicated that instructional setting tends to increase the academic achievement of higher ability students and often has a detrimental effect on lower ability learners, (Boaler, William, & Brown, 2001). In a study conducted by Ireson, Hallam, and Hurley (2001), 45 mixed secondary comprehensive schools in England were studied. The researchers found that schools set by ability encouraged a greater number of students to take science at a higher academic level, with 36% of students going on to take a science A-level compared to 25% at non-set schools. However, 17% of students from set schools were disaffected with science, stating that they never wanted to take a science class again after age 16, compared with 14% in schools with mixed-ability classes.

Despite research studies that established the efficacy of homogeneous grouping, schools have applied classification practices such as banding and setting in order to raise levels of academic achievement. While studies investigated several variations of grouping methodologies,

there continues to be an aperture in research due to the fact there have been few attempts by researchers to elucidate ways that educators can consider equality and influence practices within schools concerning grouping procedures. Trigg-Smith (2011) considered how educational policies have contributed to the decisions of educators with regard to homogeneous grouping, and how educators have considered existing theories of intelligence. Additionally, researcher discussed how student academic capability has reinforced the grouping practices and plausible frameworks for the exploration of equity within the grouping.

Negative Academic Effect of Homogeneous Grouping

In respect to the negative academic effect of homogeneous placement, Slavin (1990) conducted a meta-analysis of 29 international studies. In this study, the researcher collected data that reflected the placement of students into heterogeneous or homogeneous classrooms according to their academic capabilities, as well as student assessment scores concerning standardized test scores. Once the data were converged, Slavin found no correlation between academic achievement and ability placement. Based on their type of academic grouping, Kulik and Kulik (1987) indicated that, while higher achieving students benefited from homogeneous placement, lower performing students did not have a significant amount of improvement. While Kulik noted there was a positive effect on higher performing students, there was not a statistically significant effect in respect to the lower academically performing students as a result of homogeneous placement.

Slavin (1990) conducted one of the most comprehensive reviews of research of homogeneous grouping in elementary (1987) and secondary (1990) schools. Slavin argued that homogeneous grouping has no effects on productivity or inequality. As a result of homogeneous grouping, some of the studies indicated positive results while others indicated negative results.

Therefore, Slavin concluded that homogeneous grouping had no effect and the positive effects appeared in many studies resulted from random or systematic errors of measurement (Slavin, 1990).

In agreement with Slavin, Tieso (2003) considered homogeneous placement to be an unacceptable approach for the grouping of students. In Tieso's study, the overall impact of homogeneous placement was analyzed in respect to different academic populations. For example, students that were identified in this study were gifted, racially diverse, exceptional, and economically disadvantaged. Research studies have indicated that the utilization of homogeneously grouped classes have a negative impact upon minority students, lower socio-economic students, and students with lower academic capabilities (Tieso, 2003). Furthermore, homogeneous grouping could potentially deny students their right to an equal educational opportunity.

In the 1970's and the 1980's, several studies were published that criticized homogeneous grouping. Ethnicity and socio-economic classification of students figured prominently in the discussion (Loveless, 2013). According to Rist (1970), the researcher opined that the categorization of students by academic capability have divided students by various characteristics that are correlated statistically within areas of ability, ethnicity, and class. Critics of homogeneous grouping argued that the educational practice often separated students into socio-economic status-related groups with negative outcomes.

As one of the critics of homogeneous grouping, Rist followed a group of kindergarten students through the first years of school. Rist (1970) research study indicated that the placement of the English language art groups did not often change. He noted that the academic placement of students reflected the students' socioeconomic status. Rist (1970) opined that educators tend to

develop various expectations for groups of low and high performing students, even if those groups were given innocuous sounding names to mask their academic status.

As noted by opponents of homogeneous grouping, Collins and Gann (2013) contended that the educational practice led the lower socio-economic and minority students to low tracks where one experienced a poorer quality of academic instruction. As noted by Collins and Gan (2013), a list of the negative aspects of homogeneous grouping has been given:

- Students have a tendency to be grouped in high, middle, and low ability classes.
- Students placed in lower achieving classes can have a negative impact on one's self-efficacy.
- Often times, the students from higher socio-economic status are placed in the higher achieving classes and the students from the lower socio-economic status are placed in the lower achieving classes.

Homogeneous Grouping in Regard to High Achievers

Research has shown both advantages and disadvantages concerning grouping by capability in education. However, Catsambis and Buttaro (2012) concluded there are some benefits to homogenous classrooms grouped by ability for higher academically achieving students. Concerning kindergarten academic placement, Catsambis and Buttaro's research indicated that, when placed into small groups for reading, students in higher achieving groups had a positive attitude concerning reading. However, students placed in lower achieving groups had a negative attitude concerning reading. Catsambis and Buttaro analyzed the psycho-social impact that homogeneous-grouped classes had on students. The researchers indicated that the higher academically achieving groups had a more positive outlook on education, which correlated to their academic performance. However, the lower achieving groups had a lower

academic performance due to their negative outlook on their academics. In a similar study, Hornby, Witte, & Mitchell, 2011 found that high school students performed in like fashion. As a whole, with the exception of gifted and talented students, those grouped by ability did not perform significantly higher. With these studies in mind, it appears that homogeneous grouping increases the success of high academic achievers. However, it has a negative impact on lower academically achieving students.

Research studies have indicated that higher academically achieving students in kindergarten through grade 12 have an increased level of learning when placed in homogeneous groups (Totten & Bosco 2008). However, in the low to middle level groups, students showed higher levels of academic productivity when they were instructed in heterogeneous ability groups. Burris, Heubert, and Levin (2006) indicated contradictory results to Totten and Bosco (2008). In this study, the researchers revealed that higher academic achieving students were not affected when assimilated with students whose academic capability were lower.

Homogeneous Grouping in Mathematics and English Language Arts

To evaluate if there was a significant difference in student academic achievement scores when homogeneous grouping was implemented in elementary education, a compelling amount of research has been conducted in the subject areas of mathematics and English language arts. Even though grouping students based on English language arts ability has been historically accepted in education, Lleras and Rangel (2009) indicated that it was detrimental to academically struggling students. However, students who received instruction in a homogeneous-grouped mathematics classroom showed growth (Soloman, 2007). According to research, it was clear that there were differences between the impact of homogeneous grouping in mathematics and English language arts.

Slavin (1986) reviewed seven studies of elementary schools that homogeneously-grouped mathematic and English language art classes. According to Slavin (1986), the research studies in the best-evidence synthesis indicated that students improved more academically in regrouped classes than in heterogeneous classes, while two studies noted detrimental results. Despite this finding, in one of the studies in which students in regrouped classes did not perform better than those in heterogeneous classes (Davis & Tracy, 1963), there was no attempt made to give different and varied instructional materials to the regrouped classes. In Slavin (1992) and Kulik's (1992) analysis, the researchers noted use of the same instructional materials for all placement levels also occurred in various other studies.

Even though educators have the most significant influence on student learning, there has been a small amount of research on how instructional methods are distributed within ability-grouped classes in mathematics and English language arts (Donaldson, LeChasseur, & Mayer, 2016). The researchers examined whether teaching methodologies varied significantly among grouped levels in these core subjects. Using multilevel modeling, the researchers noted that educators of lower ability grouped classes gave less emotional and instructional support to students in their classes than did teachers of higher academically achieving classrooms. Mathematics classes reported to have higher quality instructional support for content understanding, analysis, and problem solving than English language art classes. The researchers developed case studies that illustrated how little and significant differences in instructional quality are often correlated with substantially diverging experiences for students in higher and lower ability grouped classes (Donaldson, LeChasseur, & Mayer, 2016).

Academic Effect in Mathematics

With the ample amount of studies conducted concerning the academic effect in mathematics, researchers have found that homogeneous grouping in this subject area has indicated more advanced achievement in school (Pierce, Cassady, & Adams 2011; Soloman, 2007). In situations when students were clustered, Pierce, Cassady, & Adams (2011) noted that students who participated in an academically challenging class performed significantly higher on assessments in mathematics. During this study, the researchers found that when grouping for mathematic instruction was implemented, the gifted and non-gifted students thrived academically in the classroom. Therefore, Pierce, Cassady, & Adams (2011) stated that the academically gifted students were taught by differentiated instruction with students who performed at a similar capability ability level.

In contrast to the findings of Piere, Cassady, & Adams (2011), Venkatakrishnan and William (2013) reported that grouping students based on academic capability in mathematics did not have a significant impact in respect to the high performing students. According to Venkatakrishnan and William (2013), the researchers indicated that higher academically achieving students were not advantaged significantly when homogeneously grouped. However, students progressed significantly when heterogeneously placed in mathematics. This study indicated that, when placed in mixed-ability groups, lower academically achieving students obtained the most advantage while setbacks in respect to the higher academically achieving students were minimal.

In a study conducted by Burris, Heubert, and Levin (2006), the researchers indicated contradictory results. In their study, the researchers revealed that high academically performing students were not affected when placed in classes with students whose ability was below theirs.

A longitudinal research approach was conducted which evaluated scores in standardized mathematics achievement tests in six consecutive years. The research discussed the idea that homogeneous grouping was the reason for consistently low academic achievement in public schools. The researchers considered the impact on mathematics achievement when high academically performing students were placed with students in the lower ability levels. According to Burris, Heubert, and Levin (2006), the researchers concluded that higher academically achieving students were not adversely affected by the homogeneous placement.

In contrast to Burris, Heubert, and Levin (2006), Preckel, Gotz, and Frenzel (2010) noted the effects of ability grouping in respect to mathematically gifted students on academic self-efficacy and boredom were pronounced. The researchers determined that students reported low math academic self-efficacy at an early period of the academic year. Therefore, interventions should be implemented to counterbalance this negative effect. Despite the effect on the higher academically gifted students, there is no evidence that gifted students are bored in advanced placement mathematics classes. The students gave varying reasons for the experience of boredom in class. This supports the notion that gifted classes should be provided appropriate levels of challenging tasks throughout the school year (Preckel, Gotz, & Frenzel, 2010).

In respect to homogeneous placement, one of the most noted concerns in mathematics has been the lower academically achieving groups may receive a lower quality of education overall than their higher-achieving classmates (Vogl & Preckel, 2014). With more time focused on discipline issues, the lower academically achieving students tend to receive less instructional time in contrast to the higher-achieving peers. As noted by Csikszentmihalyi & Wong (2014), educators of lower homogeneously-grouped classes are often found to be less qualified and less experienced than their peers who provide instruction to the highest achieving groups. The

researchers have indicated that teaching lower academically achieving students may be viewed as less desirable. Therefore, the teaching positions are assigned to less experienced educators. Additionally, teaching higher-academically achieving classes such as Trigonometry or Advanced Placement English language arts often requires additional training (Vogl & Preckel, 2014).

Academic Effect in English Language Arts

In the elementary school setting, teachers have utilized homogeneous placement as a way to increase student success in reading. Despite the fact that homogeneous grouping has often shown academic success for students in reading, Tach and Farkas (2006) utilized data in their research to estimate the impact of homogeneous grouping in relation to the kindergarten and first grade reading levels. The researchers indicated that prior performance on reading assessments were the most significant predictor of placement. The researchers attributed the differences in the effect in regard to socio-economics, gender, or race when homogeneous placement was first implemented. The study indicated that in kindergarten and first grade classes when homogeneous placement was introduced, a higher placement positively impacted the reading performance of students (Tach and Farkas, 2006).

In comparison to Tach and Farkas (2006), Hong, Corter, Hong, and Pelletier (2012) considered the effect that homogeneous placement in English language arts had upon kindergarten students. To ascertain how much time was spent in each group, several thousand students were observed in this research study. In contrast to Lleras and Rangel (2009), Hong et al. (2012) indicated that homogeneous placement has a positive effect on learning for low to average students when enough time was spent with various instructional methods.

Even though homogeneous placement was thought to attribute to the lack of equality in education, it has shown signs of benefitting students who speak English as a second language.

For minority students who may not regularly speak English at home, the instruction received in small reading ability groups may be advantageous (Robinson, 2008). According to this study, homogeneous placement in English language arts during kindergarten was often associated with tremendous benefits for language-minority Hispanic students when compared to others. However, if this form of placement for Hispanic students was not continued in first grade, the positive impact did not last during the summer or following year.

In 1990, principals indicated that 60% of students were homogeneously grouped in English language art classes (Loveless, 2013). However, the statistic declined to a low of 32% by 1998. In 2003, schools reported that 43% of students were grouped homogeneously in English language arts. The academic classes of history and science have less data than mathematics and English language arts. Science and history registered their highest statistical numbers in 1990 and then indicated diminished homogeneous grouping after that. Of the four academic subjects, the least amount of homogeneous grouping occurred between 1994 and 1998 (Loveless, 2013).

In California and Massachusetts, the statewide results were consistent with the national pattern of previous studies. In those two states, decline of homogeneous grouping was the most reported in the early to mid-1990's (Loveless, 2009). However, variations among the core academic subjects emerged. English language arts, science, and history heterogeneously grouped classed the most often, but educators continued to homogeneously group in mathematics. In a 2009 survey of schools in Massachusetts with eighth grade classes, 15.6% of schools offered heterogeneously-grouped mathematic classes; 49.2% offered classes with two ability levels; and 35.2% offered a variation of three levels. In other subject areas, homogeneous grouping had nearly vanished. In English language arts, 72.7% offered only heterogeneously-grouped classes, 89.8% in history, and 86.7% in science (Loveless, 2009).

Effect of Homogeneous-Grouping on Student Self-Esteem and Inequality

Relating to homogeneous placement, one concern has been the effects upon the self-esteem of students and the sense of inequality that can be associated with it. Research studies have suggested that the homogeneous placement of students in high school may not only be prejudicial, but also detrimental to the overall well-being of students. When homogeneous placement was utilized, higher academically achieving students often reported positive feelings of self-esteem, while the lower achieving students often indicated negative feelings of self-esteem (Byrne, 1988). A research study that followed sophomores to their senior year indicated that their self-efficacy remained intact for higher and middle performing students, but self-esteem dropped for the lower placed groups (Vanfossen, Jones & Spade, 1987). In mixed placement classes of English language arts and history, high school students experienced an increase of self-efficacy.

Concerning the emotional development of students, Gallagher (2011) found that many educators do not have adequate training in this area. When working with academically gifted students, teachers expressed concerns regarding elitism, as well as the need for effective socialization. In this study, Gallagher opined that educators think that homogeneous placement was a positive educational methodology; however, there was a need for heterogeneously-grouped settings that might strengthen the societal development of students, and provide a more egalitarian environment.

Powell (2008) studied whether the grouping of students in respect to English language arts capability would impact the self-efficacy of third, fourth, and fifth grade students who have below average, satisfactory, and above average reading skills. Independent t -tests indicated

significant differences in self-concept levels. Among the third graders, only the average learners significantly improved in the self-concept scores since they obtained higher scores during post-test. In regard to the 4th grade students, statistical differences existed in the self-efficacy of below average learners. Fifth grade students did not show any change in self-efficacy despite the placement (Powell, 2008).

Effect of Homogeneous-Grouping on Student Self-Esteem

Research studies have indicated that homogeneous grouping has a direct effect on student self-esteem when students received customized academic instruction (Byrne, 1988). In respect to classes that performed at a more challenging level of academics, research has shown that self-efficacy scores are higher, especially in mathematics (Reuman, 1989). Despite this research, one of the main concerns of homogeneous grouping was the negative effect that was associated with a student's self-esteem. Ireson, Hallam, and Plewis (2001) conducted a study relating to the self-esteem of students concerning the connection to their homogeneous grouping. This research studied over 3,000 students, and assessed the data based on the type of homogeneous grouping: no homogeneous grouping, limited, streaming, academic level, gender, and scores concerning one's self-esteem. According to this research study, the data indicated that students exposed to homogeneous grouping have higher self-concepts (Ireson, Hallam, and Plewis, 2001) When the overall effect of homogeneous grouping was considered, Shield (2010) indicated that students of all levels of academic capability demonstrated more self-efficacy when placed in heterogeneous classes. Despite this finding, Lou et al. (1996) found that homogeneous and heterogeneous placement had a divergent effect on the academic growth of students. When the two groups were compared, it was disclosed that the students placed in homogeneous classes accomplished more academically than the heterogeneously placed students. In contrast, Emily, Robert, and Michael

(2003) indicated that neither homogeneous nor heterogeneous placed classes are better in relating to improving the academic achievement of students.

Concerning groups performing at a higher level of academic achievement, research has shown that self-concept scores are higher, especially in mathematics (Reuman, 1989). Despite this research, one of the concerns of homogeneous grouping was the negative effect that was associated with a student's self-esteem. Ireson, Hallam, and Plewis (2001) conducted a study relating to the self-esteem of students concerning the connection to their homogeneous grouping. This research studied over 3,000 students, and assessing the data based on the type of homogeneous grouping: no homogeneous grouping, limited, streaming, academic level, gender, and scores concerning one's self-esteem. According to this research study, the data indicated that students exposed to homogeneous grouping have higher self-concept (Ireson, Hallam, and Plewis, 2001).

In another research study concerning high academic achievers, Vialle, Heaven, and Ciarrochi (2015) noted in this research study that many educators consider the correlation between student's self-efficacy and academic achievement as a confirmed fact. To disprove this commonly accepted belief, the researchers investigated the connection between self-efficacy and academic attainment in 65 higher academically achieving secondary students, and a sample was drawn from a longitudinal study of over 900 students. The researchers determined that there were no differences in measured self-efficacy between the gifted and non-gifted students.

In comparison Ireson, Hallam, and Plewis (2001), Kususanto (2010) also conducted a study regarding self-esteem and the effect of homogeneous grouping. In this research study conducted in Malaysia, 302 students from four public high schools completed a survey on the topic of self-esteem. This study revealed a noteworthy difference in teacher perception and self-

esteem while also showing academically higher achieving students scored considerably higher than academically lower achieving students. Students in lower academically achieving classes indicated that their instructors were domineering; however, higher academically achieving students indicated that their teachers were encouraging. In conclusion, Kususanto indicated in this research study that there was a positive relationship between the teacher's conception of students and the student's self-esteem. For this reason, teachers might have different impacts on the students when considering their academic achievement level. Therefore, it might inadvertently affect the self-efficacy of students.

Effect of Homogeneous-Grouping on Social Inequality

When discussing the impact of homogeneous grouping and the effect on social inequality, Kintz (2011) opined that classifying students into groups concerning one's academic achievement was not a normal societal procedure. Researchers have concurred that categorizing students into homogeneous groups tended to escalate societal injustice. Fendler and Muzaffar (2008) composed an article in which the researchers considered the history of the bell curve. In educational settings when a true bell curve was utilized to calculate assessments scores or the classifying of students, there will often be a component of insufficiency. In respect to the history of education, the bell curve assured that many students are "average" while the other students are either advanced or failing. Fendler and Muzaffar (2008) indicated that the objective of education is to help students achieve academically. For this reason, when students are classified based on their academic ability level, there was one of two formulations: social alteration or social competition.

Ireson and Hallam (2001) indicated that there was a significant impact on social comparison that students experience in relating to homogeneous grouping. The review of

literature has indicated there are both negative effects of homogeneous-grouping, and no correlation at all between self-esteem scoring and students' views of their capabilities and self-confidence in mathematics and science (Ireson & Hallam, 2001). However, this research study indicated that there was a statistically significant difference between negative self-esteem scores and homogeneous-grouping in the subject area of English language arts (Chorzempa and Graham, 2006)

In agreement with Ireson and Hallam (2001), Kintz (2011) also determined a disparity existed between students in high and low homogeneous groups. Kintz revealed students from higher socio-economic backgrounds tended to be in higher academic ability groups. His study also suggested when students were tracked at an early age they may be placed into a less rigorous educational path. The study concluded homogeneous-grouping negatively impacts lower academically grouped students for a variety of reasons including less rigorous instruction being implemented and friendships formed with less motivated students.

According to Flashman (2012), the academic achievement of students was often impacted by acquaintanceships. In Flashman's research, a school system was studied with two diverse settings, one school system was rural, and the other was located in an urban area. The process for evaluating the school systems was identical. Despite the diversity of the locations, Flashman's study indicated that higher academically performing students were friends with other high achievers. The lower placed students were friends with other lower performing students. However, in the rural school setting, academic achievement level did not have a significant impact on the friendships. Flashman indicated this might be due to the smaller numbers of students. In conclusion, Flashman (2012) opined that acquaintances are often formed by students with similar backgrounds, which could have an impact on one's future success.

Despite the fact that homogeneous placement allows educators to meet the academic needs of a large group of students in an instructional setting, research has shown that there are drawbacks to this type of grouping. Ireson and Hallam (2001) conducted a study comparing self-concept scores of students in homogeneously-grouped schools and heterogeneously-grouped schools. These researchers found that students who have experienced the instructional setting not grouped by ability have a higher self-concept with regard to school.

In comparison to Ireson and Hallam (2001), Neihart (2007) collected research that focused on the socio-affective impact that homogeneous grouping has on academically gifted students. Neihart's research revealed that, while gifted students benefited from homogeneous grouping, some students benefited even more from interacting and socializing with peers of different abilities. Neihart's study further concluded that students prematurely accelerated to an inappropriate level might not benefit emotionally from social interaction the same as those who are academically gifted or those who perform at a lower academic ability. Because Neihart's study indicated some gifted students often function better in a heterogeneous setting, his study concluded an individual assessment must determine if a gifted student functions better socially, emotionally, and academically in a homogenous setting.

Standardized Assessment and Learning

When analyzing academic assessments and its impact on student learning, many components must be taken into consideration. During the course of a student's academic career, students will participate in many forms of assessment. Pekrun, Goetz, Titz, and Perry (2002) indicated that one's emotions have an impact on students' academic achievement, learning, and motivation. Likewise, Pekrun et al. (2002) discovered that students reported having both positive and negative emotions toward school and their overall academic experience. The most

commonly reported academic emotion was anxiety. Due to the high levels of anxiety that have been reported, it should be important to analyze student learning and discover the most appropriate method of assessment for individuals (Pekrun et al., 2002).

Since the Progressive period (1890-1920), standardized assessment or high-stakes testing has been used in education in the United States. During this period, the United States focused on industry and tracking students in order to meet the needs of the current economy. During its origin, the standardized intelligence tests were created by educators to place students into tracks according to their career outlook (Emery, 2007). Through the years, standardized testing has developed into high-stakes testing, which was influenced by corporate-inspired educational reform (Emery, 2007).

For over a century, standardized testing for assessing student's aptitudes and academic achievement has continued to have an increasingly prominent role in shaping the American educational system (Nichols & Berliner, 2007). In the United States, the well-known tests used throughout the elementary and secondary grades are the Iowa Test of Basic Skills (ITBS), TerraNova, Stanford Achievement Tests (SAT), and the Metropolitan Achievement Test (MAT). In addition, many secondary students have prepared for college entrance by taking the SAT exam administered by the College Board or the ACT exam administered by the American College Testing Program. If one is applying for entrance to graduate school, one is often required to take other exams, usually the Graduate Record Examination (GRE) and/or the graduate tests for law (LSAT), business (GMAT), or medicine (MCAT) (Nichols & Berliner, 2007).

Over the past thirty years, there has been an initiative to standardize education as well as assessments, in order to maintain high academic standards in the United States. During the 1980s, President Ronald Reagan implemented three goals of reform to address the perceived

lackluster performance of the United States educational system. The first goal focused on increasing graduation requirements, standardizing the curriculum, and increasing testing and certification requirements for educators. Focusing on the level of education on a local level, the second goal proposed an increase in professionalism and empowerment of teachers, while the third and final goal concentrated on students and their learning (Bohrnstedt, 2013). During this time, Emory (2007) contended top CEOs of the country agreed to focus on education reform through high-stakes testing because there was an increased involvement of the Federal government to standardize curriculum and assessment.

During his administration, President Clinton initiated education reform through Goals 2000, which focused on accountability for the academic achievement of students, high-stakes testing, and state content standards. As a result of President George W. Bush's No Child Left Behind Act, teacher accountability, student assessment, and a push for increased rigorous standards were further accentuated. Unfortunately, this reform coincided with a period of widening of the achievement gap where higher achieving students were making greater academic strides, but struggling students were less successful. Recently, President Obama's administration has focused on student growth achievement rather than measuring student academic achievement. In addition, since 2008 student achievement measurement has been linked to legislation in many states that links teacher evaluation with student growth. (McGuinn, 2011).

Measuring Student Learning

When discussing homogeneous placement and its academic impact, Sackett, Borneman, and Connelly (2008) found there are some standardized assessments that accurately measured student learning. However, in Finland, it is theorized that measuring student learning does not have to be assessed through high-stakes testing (Sahlberg, 2007). The students in this

Scandinavian country do not take any standardized assessments until they finish upper-secondary school. In Finland, for the most part, primary school students are not tested. Primary school is utilized for sustaining natural curiosity. Teachers are allowed freedom in curriculum planning (McCarter, 2014; Sahlberg, 2007). Through this freedom, educators are allowed to focus their attention on providing students with basic knowledge in the first 6 years and move into reasoning and problem solving in the later years of study. Within the Finnish educational school system, educators have adopted the following principles that outline their beliefs regarding education:

- The majority of teachers in Finland have a master's degree or higher.
- The schools in Finland have a "culture of trust" between students, parents, teachers, and administrators.
- The leadership in Finland is supportive of students and teachers.

Even though Finland does not focus on standardized assessments, the country has increased student academic accountability through other international assessments, such as The International Mathematics and Science Study (TIMSS), and the Programme for International Student Assessment (PISA) (McCarter, 2014; Sahlberg, 2007). While Finland has not implemented standardized assessments within its curriculum, it has seen increased student success based on these two assessments and an increase in students who complete tertiary education (Sahlberg, 2007)

In agreement with Sahlberg (2007), Popham (2011) indicated that outcomes differ as to the precision of standardized assessments, in relation to student learning. Popham noted that there are few standardized assessments that do an adequate job in translating student learning in relation to test scores. Even though standardized assessments do not always indicate accurate reflections of student learning, Popham indicated that teachers must be assessment literate.

Popham suggested an evaluation method that produces a more accurate reflection of learning should replace a standardized assessment to monitor school and academic growth success, as well as accountability for funding (McCarter, 2014).

In education there are many concerns in regard to the results of standardized assessments and how effectively it measures student learning. Sackett, Borneman, and Connelly (2008) found that there are common beliefs regarding assessments. While these beliefs are widely accepted, there has been research that contradicted these findings relating to each of the following statements: (Sackett, Borneman, and Connelly, 2008; McCarter, 2014):

- Standardized assessments are often poor predictors of student academic achievement.
- Standardized assessments often measure short-term learning.
- Standardized assessments are often used for predictive value.
- Standardized assessments often reflect socio-economic status rather than developed capabilities of students.
- Standardized assessments are often biased against minority groups.

In contrast to the study conducted by Popham (2011), after analyzing data from previous studies, Sackett et al. concluded that, if tests are constructed correctly, they could accurately portray student learning. Sackett, Borneman, and Connelly (2008) found the following results:

- High-stakes tests are generally valid regarding the correlation between student ability and test results.
- The validity of the standardized test is not a result of the socio-economic status of students.
- If the teachers coach their students, it is not an indicating factor of performance on the results of the test.

- High-stakes tests are not biased against minority students.
- If teachers try to motivate their students prior to the test, it does not always have a positive impact on their performance.

According to this study, Sackett et al. (2008) indicated that there was a positive correlation between student ability and standardized test scores; however, they cautioned that not all standardized tests are capable of providing an accurate measure of student learning. The researchers stated that some tests that claim to be valid could be used inappropriately, which causes them to lose their validity.

Limitations of Standardized Assessments

In respect to the limitations of standardized assessments, Lee (2008) indicated that the high-stakes testing policy implemented by the government had zero effect on the academic achievement gap. The researcher noted there was still a large racial academic achievement gap that continues to grow. Lee (2008) concluded that the evidence indicating that high-stakes testing is reliable is limited. He indicated that further study needs to be conducted in order to determine the extent to which high-stakes testing policies impacted student achievement.

In comparison to Lee (2008), Popham (2007) discussed the implications of an insensitive standardized assessment against an instructionally sensitive test. Instructionally insensitive tests, such as standardized assessments, do not allow for interpretation of scores as it pertains to either strong or weak instruction. When an assessment was considered sensitive, Popham (2007) stated that it could be used to measure quality instruction. However, Popham found that current high-stakes tests do not meet the criteria to be considered sensitive. Therefore, they are considered insensitive and a poor measure of teacher instruction and student learning.

The implications of No Child Left Behind (NCLB) were studied by Guisbond and Neill (2004). This study indicated two main assumptions relative to this legislation. The assumption that boosting of test scores being a primary goal for schools assumes education can have a one-size-fits-all approach. Differentiating instruction to meet the needs of the students and consideration of social goals of schools would not be considered with this approach (Guisbond & Neill, 2004). Furthermore, poor teaching resulted in low student performance that can be rectified by sanctions. Even though schools do not receive Adequate Yearly Progress (AYP) any longer, 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. 199).

Furthermore, Guisbond and Neill (2004) indicated the federal, state, and local governments should work together to offer a rich curriculum learning opportunity to all students. Additionally, educators should use a variety of assessments to evaluate student learning rather than a single standardized assessment. The accountability of learning should not be placed on the teachers only. Additionally, the researchers noted that the entire school system, parents, students and members of the community should accept their role in the learning process.

In agreement with Lee (2008) and Popham (2007), Guisbond and Neill (2004) indicated that student learning should be evaluated during the school year through formal and informal evaluations and standardized assessments should be reduced or completely eliminated. Furthermore, the researchers stated that if these principles were enacted into the legislative methods in education, students would have the opportunity to learn and not be limited based on the scores they earned on one standardized assessment.

Summary

In summary the research literature in respect to homogeneous placement has been conflicting. Over the years homogeneous grouping has been a controversial educational practice (Ansalone & Biafora, 2004; Rubin, 2008). Due to the fact that it, "has been the subject of more research studies, well over 500, than almost any other educational practice has been a controversial topic" (George & Alexander, 2003, p. 414). The proponents of homogeneous grouping have maintained that it specifically focuses instruction on the needs of groups of students. However, opponents of homogeneous grouping have maintained that the expected advantages are often not materialized. According to Snider and Schumitsch (2006), homogeneous grouping has promoted a negative stigma and destroyed academic motivation, especially among the lower ability students.

Homogeneous grouping in the academic setting has been the subject of debate for several years (Hopkins, 2008). The topic has been controversial due to the fear that it will hinder the advancement of students with lower academic capability since there will be an emphasis on basic knowledge instead of advanced learning. On the other hand proponents have insisted that grouping by academic capability has the potential of improving student achievement by increasing the levels of motivation. The only certain conclusion is that further research on ability grouping will benefit students, educators, and school administrators.

CHAPTER THREE

Methods and Procedures

The objective of this study was to ascertain if there was a statistically significant difference between school-wide student academic achievement in grades three, four, and five, based on the type of grouping in classroom settings. This chapter focused on the methodology that was used in this non-experimental, quantitative study to evaluate if there was a statistically significant difference in the scores of students in proficiency categories on the Tennessee Comprehensive Assessment Program (TCAP) in mathematics and English language arts, between homogenously-grouped classrooms and classrooms that were not homogeneously grouped. This study focused on the academic achievement proficiency scores on the Tennessee Comprehensive Assessment Program (TCAP), as indicated by the Tennessee Department of Education website and the Pearson published reports for schools in a rural county in East Tennessee. The results were limited to grades three, four, and five in the subject areas of mathematics and English language arts for the 2014-2015 school year. Comparisons were made in the subject areas of mathematics and English language arts, as well in grades three, four, and five.

Research Questions

The research study was guided by the following research questions.

Research Question 1

Is there a significant difference in the scores of students in each of the Tennessee Comprehensive Assessment Program (TCAP) levels: Proficient and Below Proficient for grades three, four, and

five in mathematics between schools that group students homogeneously and schools that do not?

Research Question 2

Is there a significant difference in the scores of students in each of the Tennessee Comprehensive Assessment Program (TCAP) levels: Proficient and Below Proficient for grades three, four, and five in English language arts between schools that group students homogeneously and schools that do not?

This chapter described the population and sample, and provided a description of the instruments used for this study. Research procedures were also outlined. Finally, data analysis methods were identified as a means to answer the research questions.

Instrumentation

The elementary school data were obtained from the Tennessee Department of Education website as well as the Pearson published reports. The state of Tennessee reports student scores by the following: Advanced, Proficient, Basic, and Below Basic. As noted by the Tennessee Department of Education website, the percentage of scores required to be "proficient" are between 30% to 60%. However, for this research study, the achievement levels were identified as Proficient and Below Proficient. For the purposes of this study, Basic and Below Basic were combined and referred to as below proficient. In addition, Proficient and Advanced were combined and referred to Proficient. McMillan and Schumacher (2011) noted that standardized assessments are reliable, and the results are objective. The Tennessee Comprehensive Assessment Program (TCAP) was described by the Tennessee Department of Education as the following: it is a multiple choice assessment that evaluates student skills in reading, language arts, mathematics, science, and social studies (TDOE, 2014). However, for this research study,

the achievement levels were identified as Proficient and Below Proficient. McMillan and Schumacher (2011) noted that standardized assessments are reliable, and the results are objective. The Tennessee Comprehensive Assessment Program (TCAP) was described by the Tennessee Department of Education as the following: it is a multiple choice assessment that evaluates student skills in reading, language arts, mathematics, science, and social studies (TDOE, 2014).

Research Participants and Setting for the Study

Population. In order to control for cultural, economic, and societal variations, the population for this research study was identified as public elementary schools in a rural county in East Tennessee. Within this research study, eight public elementary schools participated. Stratified sampling was applied because not all elementary schools fall into the classification of either homogeneously-grouped or non-homogeneously grouped classroom settings. The sample size was based on the number of schools that implemented this type of placement for mathematics and English language arts.

Setting. The setting was defined as public elementary schools located in a rural county in upper East Tennessee. The total student enrollment in grades three, four, and five in this rural county is 1,654. Of this total number of students, 1,650 students were included as part of this study. The discrepancy in the number of students enrolled and took the test was assumed by the researcher to be due to absence or parental request to not take the assessment.

Data Collection Procedures

The research data were collected from the Tennessee Department of Education School Report Cards website as well as the Pearson Published Reports. To assure reliability and validity within the research study, criterion-referenced assessment scores were utilized. According to

McMillan and Schumacher (2010), standardized assessments provide scoring consistency. Appendix A includes a copy of the researcher's request to conduct the study in the selected school system. A letter of consent was given by the director of schools for the participating system. Appendix B provides a sample of the director's letter. When collecting the research data, concerning the classification of homogeneous grouping implemented, the researcher contacted the administrator at each elementary school to assure that there was clear communication. Appendix C provides a copy of the initial email to participants. To code the schools as either homogeneously-grouped or heterogeneously-grouped classrooms, the researcher contacted each principal by school email. A copy of the questions for participants is found in Appendix D. In the email to the school administrator, the researcher defined both of these classifications. The school administrator was asked to code the grade levels for three, four, and five and for the subject areas of mathematics and English language arts as either homogeneously-grouped classrooms or classrooms heterogeneously-grouped for the 2014-2015 school year. The data collected for this research study were placed into a two-sample z-test and steps were followed as outlined by Green and Salkind (2011). Once the elementary schools were anonymously coded, the researcher assessed the Tennessee Comprehensive Assessment Program data located on the Tennessee Department of Education website as well as the Pearson Published Reports.

Validity, Reliability, and Bias

In an attempt to avoid bias within the research study, the researcher took every precaution to assure that the study was reliable and valid. When devising the research study, four types of design validity were addressed: construct validity, external validity, internal validity, and statistical conclusion validity (McMillan & Schumacher, 2010). To avoid coding mistakes, it was imperative to ensure that correct data were collected concerning information from school

administrators. During the data analysis, committee members were conferred with in order to ensure that the correct tests were run and that the data were read accurately. Once the data were analyzed, the research data were cross-checked in order to ensure that it was entered correctly, which produced a coefficient of agreement (McMillan & Schumacher, 2010). In order to avoid bias, the researcher asked for a review from committee members.

Human Safety

In order to assure human safety as part of this research study, the Institutional Review Board evaluated the application before the research was conducted. The Institutional Review Board determined that this was not a research concerning human subjects. Therefore, the safety of human subjects was ensured.

Ethical Issues

When considering the ethical issues of this research study, transparency was imperative (McCarter, 2014; McMillan & Schumacher, 2010). In order to provide all information, procedures and research data were reported in Chapters Four and Five. Due to the fact that the research data gathered for this study was considered public information, privacy and confidentiality issues were not a concern. Furthermore, taking action to avoid researcher bias was essential in the conduction of this research study.

Data Analysis Procedures

The data were entered into a spreadsheet (McCarter, 2014; McMillan & Schumacher, 2010). Two-sample z-tests were run in order to determine the connection between school academic achievement scores and the type of homogeneous placement. To determine if there was a statistical connection in scores between homogeneous grouping and heterogeneous grouping

among grades three, four, and five, in both mathematics and English language arts, two-sample z-tests were run.

Summary

The objective of this research study was to ascertain if there is a statistically significant difference between elementary school student academic achievement scores in grades three, four, and five based on the identified placement: homogeneously-grouped or heterogeneously-grouped. The school-wide Tennessee Comprehensive Assessment Program (TCAP) standardized assessment data were collected from the Tennessee State Department of Education website for the 2014-2015 school year for grades three, four, and five, in mathematics and English language arts. The elementary school administrators from a rural county in East Tennessee were contacted by the researcher to ascertain if students were grouped homogeneously for mathematics and English language arts, or if they were placed in heterogeneous classrooms for instruction. Two-sample z-tests were conducted to determine the significance between academic achievement and the type of grouping in grades three, four, and five in mathematics and English language arts.

CHAPTER 4

Analysis of Data

The objective of this research study was to ascertain if there was a statistically significant difference between student academic achievement scores in grades three, four, and five according to the type of placement in instructional settings, homogeneously grouped or heterogeneously grouped. The Tennessee Comprehensive Assessment Program (TCAP) assessments were divided by the subject areas of mathematics and English language arts. The scores of students were assessed in respect to the proficiency levels of proficient and below proficient. To determine if there was a statistically significant difference between schools that group students homogeneously and schools that do not, a two-sample z-test was conducted. The data were collected from elementary schools from a rural county in East Tennessee. The Tennessee Comprehensive Assessment Program (TCAP) scores were analyzed for grades three, four, and five in the subject areas of English language arts and mathematics from the 2014-2015 school year. The Tennessee Comprehensive Assessment (TCAP) data were retrieved from the Tennessee Department of Education website as well as the Pearson published reports. For this chapter, the data were analyzed in regards to two research questions and four null hypotheses.

Analysis of Research Questions

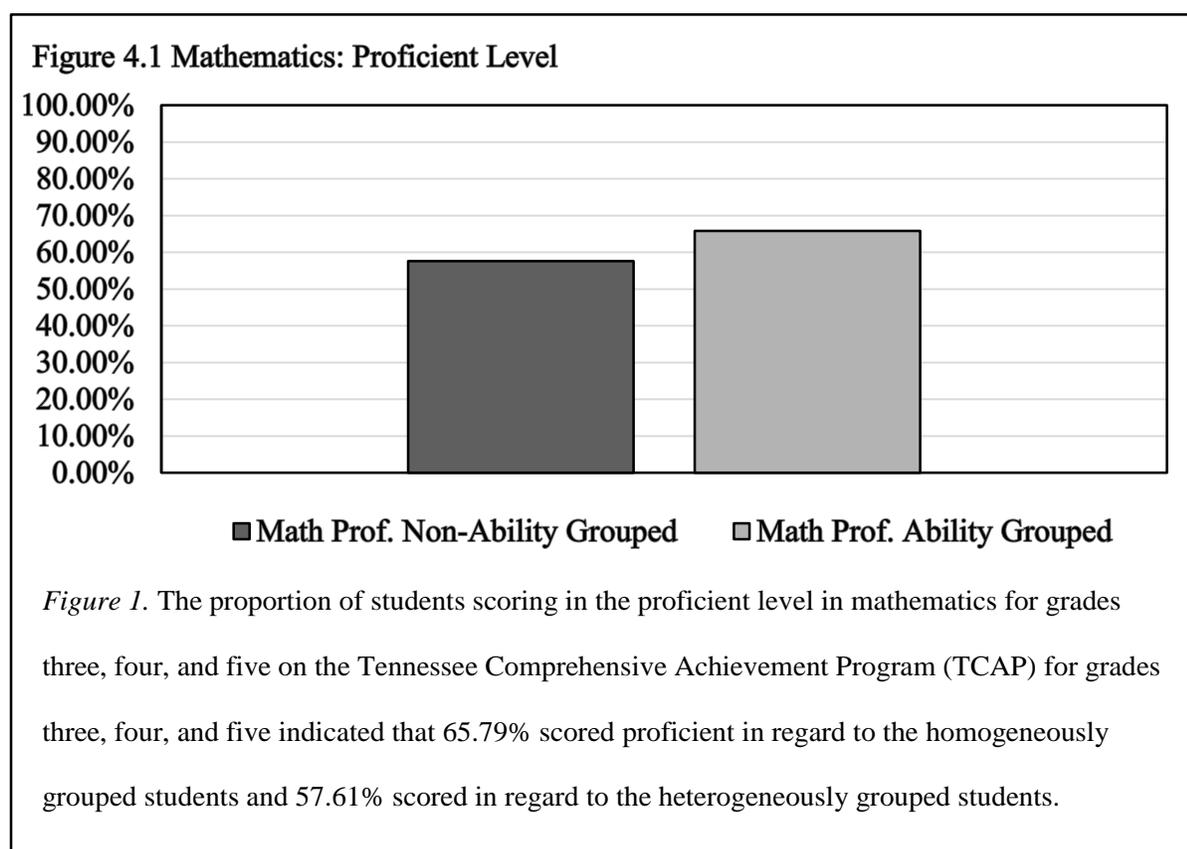
Research Question 1. Is there a significant difference in the scores of students in the Tennessee Comprehensive Assessment Program (TCAP) levels: proficient and below proficient for grades three, four, and five in mathematics between schools that group students homogeneously and schools that do not?

Ho1: There is no significant difference in the proportion of students in respect to the Tennessee Comprehensive Achievement Program (TCAP) in the proficient level for grades three, four, and five in mathematics between schools that homogeneously group and schools that heterogeneously group students by academic capability.

To determine if the mean proportion in mathematics for grades three, four, and five, among groups of students who scored in the proficient category on the 2014-2015 Tennessee Comprehensive Achievement Program (TCAP) varied among those who received instruction in a homogeneously grouped environment and those who received instruction in a heterogeneously grouped environment, a two-sample z-test was conducted. The total number of students that were included in this research study was 1,650. The students who scored in the proficient category varied significantly among those who received instruction in a homogeneously grouped setting in comparison to those who received instruction in a heterogeneously grouped setting. The proportion of students who scored in the category was the test variable and the grouping variable was the type of grouping: homogeneously grouped or heterogeneously grouped. The z-test showed a z score of 2.2, which falls to the positive side within the range of -1.96 to 1.96. There was a significant effect for those students who were homogeneously grouped, $z(2.2)$ and $P=.0276$. For this reason, the first null hypothesis was rejected. The proportion of students who scored in the proficient category in homogeneously grouped classrooms ($M = 31.25, SD = 17.32$) was not comparable to those who were in a heterogeneously grouped classroom ($M = 17.80, SD = 11.57$).

The Tennessee Comprehensive Assessment Program (TCAP) are standards-based, criterion-referenced tests. Students are measured by how well they have mastered grade-specific skills, instead of being compared to the performance of other test takers. The Tennessee

Comprehensive Assessment Program (TCAP) tests report students' performance in the subject areas of English language arts, mathematics, science, and social studies. For accountability purposes, the state of Tennessee has defined four academic achievement levels. The academic achievement levels are Below Proficient, Proficient, Above Proficient, and Advanced. However, for this research study, the achievement levels were identified as Proficient and Below Proficient. The student's performance on the Tennessee Comprehensive Assessment Program in mathematics and English language arts determines one's achievement level in these areas. Figure 4.1 indicates the proportion of students who scored in the proficient level in mathematics for grades three, four, and five.



Ho2: There is no significant difference in the proportion of students in the Tennessee Comprehensive Achievement Program (TCAP) in the below proficient level for grades three,

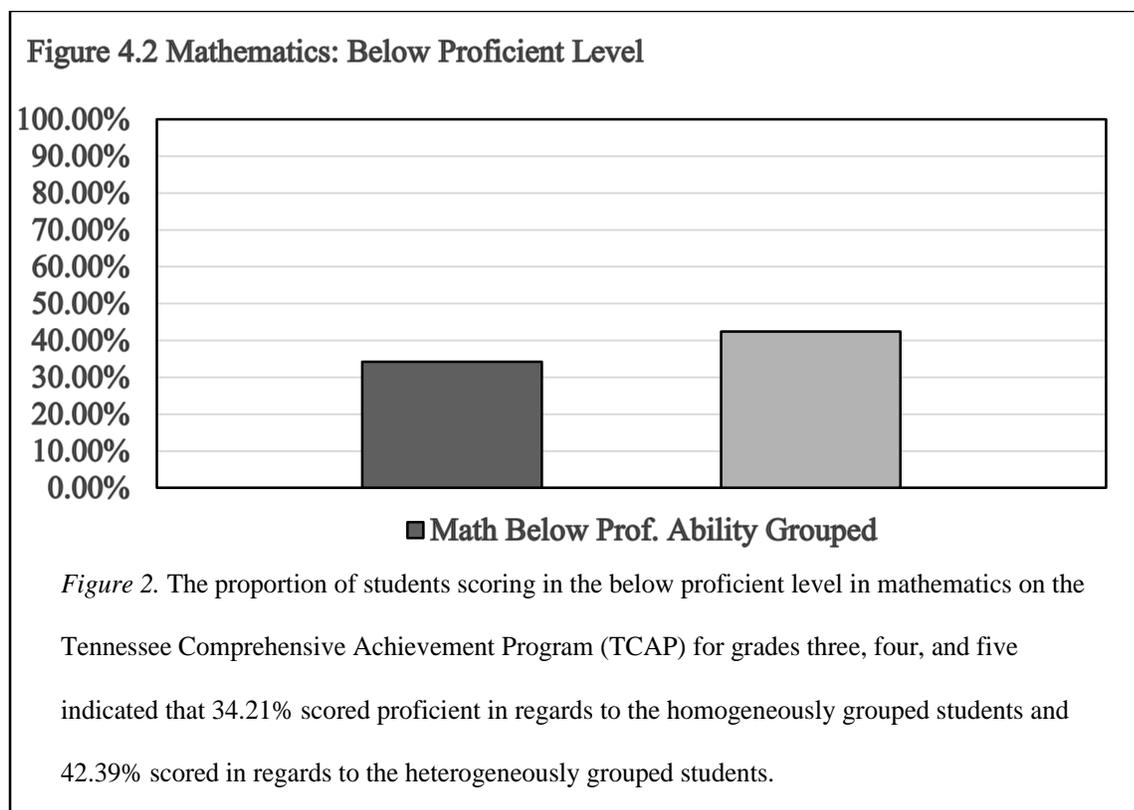
four, and five in mathematics between schools that homogeneously group and schools that heterogeneously group students by academic capability.

To determine if the mean proportion in mathematics, for grades three, four, and five among students who scored in the below proficient category varied among those who received instruction in a homogeneously grouped environment and those students who received instruction in a heterogeneously grouped environment, a two-sample z-test test was conducted. The students who scored in the below proficient category varied significantly among those who received instruction in a homogeneously grouped academic setting in comparison to those who received instruction in a heterogeneously grouped setting. The proportion of students who scored in the category was the test variable, and the grouping variable was the type of grouping: homogeneously grouped or non-homogeneously grouped. The z-test showed a z score of -2.2 which falls to the negative side within the range of -1.96 to 1.96. There was a significant effect for those students who were homogeneously grouped, $z (-2.2)$ and $p = .0276$. For this reason, the second null hypothesis was rejected. The proportion of students who scored in the below proficient category in homogeneously grouped classrooms ($M = 16.25$, $SD = 9.26$) was not comparable to those who were in a heterogeneously grouped classroom ($M = 13.10$, $SD = 8.84$).

In Figure 4.2, the proportion of students who scored in the below proficient level in mathematics for grades three, four, and five are noted. The data indicated that 34.21% of students scored proficient in respect to the homogeneously grouped students and 42.39% scored in respect to the heterogeneously grouped students on the Tennessee Comprehensive Achievement Program for the 2014-2105 school year. The state of Tennessee reports student scores by the following: Advanced, Proficient, Basic, and Below Basic. However, for this research study, the achievement levels were identified as Proficient and Below Proficient. The

Tennessee Comprehensive Assessment Program (TCAP) was described by the Tennessee Department of Education as the following: it is a multiple choice assessment that evaluates student skills in reading, language arts, mathematics, science, and social studies (TDOE, 2014).

Figure 4.2 indicates the proportion of students who scored in the below proficient level in mathematics for grades three, four, and five.



Research Question 2. Is there a significant difference in the scores of students in each of the Tennessee Comprehensive Assessment Program (TCAP) levels: Proficient and Below Proficient for grades three, four, and five in English language arts between schools that group students homogeneously and schools that do not?

Ho3: There is no significant difference in the proportion of students in the Tennessee Comprehensive Achievement Program (TCAP) in the proficient level for grades three, four, and

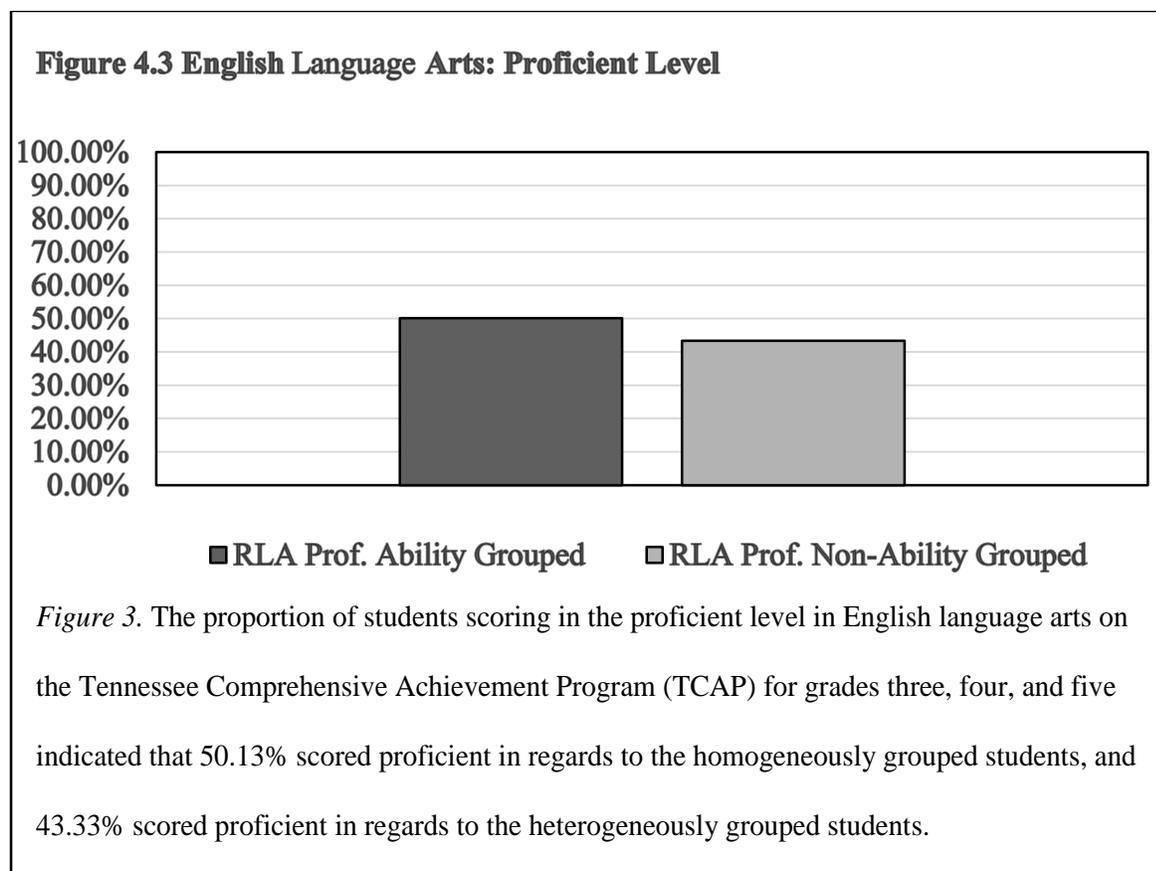
five in English language arts between schools that homogeneously group and schools that heterogeneously group students by academic capability.

To determine if the mean proportion in English language arts for grades three, four, and five among students who scored in the proficient category varied among those who received instruction in a homogeneously grouped environment and those students who received instruction in a heterogeneously grouped environment, a two-sample z-test was conducted. The students who scored in the proficient category varied among those who received instruction in a homogeneously grouped academic setting and those who received instruction in a heterogeneously grouped setting. The proportion of students who scored in the category was the test variable, and the grouping variable was the type of grouping: homogeneously grouped or heterogeneously grouped. The z-test showed a z score of 1.81, which falls to the positive side within the range of -1.96 to 1.96. There was not a significant effect for those students who were homogeneously grouped, $z(1.81) =$ and $p = .0701$. For this reason, the third null hypothesis failed to be rejected. The proportion of students who scored in the proficient category in homogeneously grouped classrooms ($M = 23.88$, $SD = 12.08$) was comparable to those who were in a heterogeneously grouped classroom ($M = 11.92$, $SD = 9.27$).

In Figure 4.3, the proportion of students who scored in the proficient level in English language arts for grades three, four, and five are noted. The data indicated that 34.21% of students scored proficient in respect to the homogeneously grouped students and 42.39% scored in respect to the heterogeneously grouped students on the Tennessee Comprehensive Achievement Program for the 2014-2105 school year. The state of Tennessee reports student scores by the following: Advanced, Proficient, Basic, and Below Basic. However, for this research study, the achievement levels were identified as Proficient and Below Proficient. The

Tennessee Comprehensive Assessment Program (TCAP) was described by the Tennessee Department of Education as the following: it is a multiple choice assessment that evaluates student skills in reading, language arts, mathematics, science, and social studies (TDOE, 2014).

Figure 4.3 indicates the proportion of students who scored in the proficient level in English language arts for grades three, four, and five.



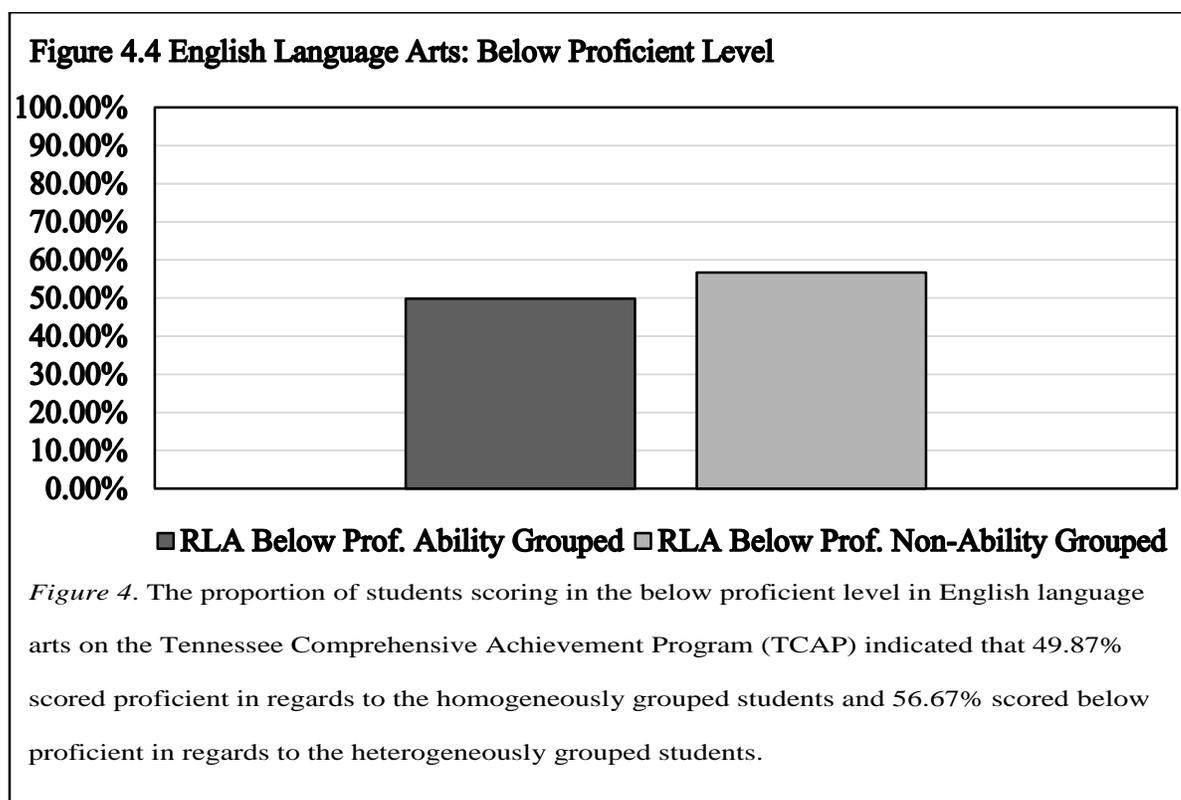
Ho4: There is no significant difference in the proportion of students in the Tennessee Comprehensive Achievement Program (TCAP) in the below proficient level for grades three, four, and five in English language arts between schools that homogeneously group and schools that heterogeneously group students by academic capability.

To determine if the mean proportion in English language arts for grades three, four, and five among students who scored in the below proficient category varied among those who

received instruction in a homogeneously grouped environment and those students who received instruction in a heterogeneously grouped, a two-sample z -test was conducted. The students who scored in the below proficient category did not vary significantly among those who received instruction in a homogeneously grouped academic setting in comparison to those who received instruction in a heterogeneously grouped setting. The proportion of students who scored in the category was the test variable, and the grouping variable was the type of grouping: homogeneously grouped or heterogeneously grouped. The z -test showed a z score of -1.81 , which fell to the negative side within the range of -1.96 to 1.96 . There was not a significant effect for those students who were homogeneously grouped, $z (-1.81)$ and $p = .0701$. For this reason, the fourth null hypothesis failed to be rejected. The proportion of students who scored in the below proficient category in homogeneously grouped classrooms ($M = 23.75$, $SD = 12.71$) was comparable to those who were in a heterogeneously grouped classroom ($M = 15.58$, $SD = 11.91$).

In Figure 4.4, the proportion of students who scored in the proficient level in English language arts for grades three, four, and five are noted. The data indicated that 34.21% of students scored proficient in respect to the homogeneously grouped students and 42.39% scored in respect to the heterogeneously grouped students on the Tennessee Comprehensive Achievement Program for the 2014-2105 school year. The state of Tennessee reports student scores by the following: Advanced, Proficient, Basic, and Below Basic. However, for this research study, the achievement levels were identified as Proficient and Below Proficient. The Tennessee Comprehensive Assessment Program (TCAP) was described by the Tennessee Department of Education as the following: it is a multiple choice assessment that evaluates student skills in reading, language arts, mathematics, science, and social studies (TDOE, 2014).

Figure 4.4 indicates the proportion of students who scored in the below proficient level in English language arts for grades three, four, and five.



Summary

In chapter four, the 2014-2015 Tennessee Comprehensive Assessment Program (TCAP) scores for elementary school students in the research county were reported for grades three, four, and five in the achievement levels of Proficient and Below Proficient. The elementary school principals were emailed, and the data were sorted in regards to the type of grouping implemented: homogeneously grouped or heterogeneously grouped. In the research county, there are eleven schools with grades three, four, and five. Of the eight schools that participated in this research study, three homogeneously grouped by academic capability, and five heterogeneously grouped.

In this research study, there were two research questions and four null hypotheses. The research data were obtained from the Tennessee Department of Education state report card website, Pearson published reports, and individual school administrators. A sequence of two-sample z-test was conducted to ascertain if there was a significant difference among the group of student scores, in regards to the reporting categories and the type of grouping. The results of this research study indicated that there was a significant difference in proportions of students in two of the four categories, in respect to the type of grouping implemented. Therefore, two of the four null hypotheses were rejected. The null hypotheses that were rejected were Ho1 and Ho2.

CHAPTER 5

Conclusions, Implications, and Recommendations

In the United States, placing students into ability groups based on their academic strengths and weaknesses has been a common practice (Davidson, 2009; Slavin, 1987). Homogeneous grouping and its modifications have been among the primary educational practices of public school systems in the United States. In the field of education, homogeneous grouping has been practiced in the United States for over a hundred years (Rubin, 2008). In this study, the research indicated both detrimental and positive results in respect to homogeneous grouping in education.

Summary of Results

The objective of this research study was to ascertain if there was a statistically significant difference between student academic achievement scores in grades three, four, and five according to the type of placement in instructional settings, homogeneously grouped or heterogeneously grouped. The Tennessee Comprehensive Assessment Program (TCAP) scores of students concerning the proficiency levels, Proficient and Below Proficient categories, were grouped by grade level and subject area: mathematics and English language arts. To determine if there was a statistically significant difference between schools that group students homogeneously and schools that were heterogeneously grouped, a two-sample z-test was conducted. The data were collected from elementary schools from a rural county in East Tennessee. The Tennessee Comprehensive Assessment Program (TCAP) scores were analyzed for grades three, four, and five in the subject areas of mathematics and English language arts

from the 2014-2015 school year. The Tennessee Comprehensive Assessment (TCAP) data were retrieved from the Tennessee Department of Education website and the Pearson published reports. The data were analyzed in regard to two research questions and four null hypotheses.

Research Question 1

Is there a significant difference in the scores of students in each of the Tennessee Comprehensive Assessment Program (TCAP) levels: Proficient and Below Proficient for grades three, four, and five in mathematics between schools that group students homogeneously and schools that do not?

A two-sample z-test was conducted to determine if there was a significant difference in the score of students in the proficiency levels for grades three, four, and five in mathematics between elementary schools that homogeneously group and those that do not. The test results indicated there was a significant difference between students who received instruction in homogeneously grouped classrooms and those that did not.

Research Question 2

Is there a significant difference in the scores of students in each of the Tennessee Comprehensive Assessment Program (TCAP) levels: Proficient and Below Proficient for grades three, four, and five in English language arts between schools that group students homogeneously and schools that do not?

A two-sample z-test was conducted to determine if there was a significant difference in the score of students in each of the proficiency levels for grades three, four, and five in English language arts between elementary schools that homogeneously group and those that do not. The test results indicated there was not a significant difference between students who received instruction in homogeneously grouped classrooms and those that did not.

Implications for Practice

In respect to the literature on the topic of homogenous grouping, the research indicated varied results as to the impact on students' academic achievement (Reuman, 1989). In general, research on the topic revealed that higher achievers often had academic gains when placed in homogeneously grouped classes. However, the middle to low achieving students did not demonstrate academic gains (Plunkett & Kronborg, 2011). Research has indicated that homogenous grouping does not have a profound impact on the self-efficacy of elementary students and one's conception of learning. Despite this information, researchers have noted that middle and high school students often have lower self-efficacy scores, as well as a negative perception of school (Hong et al, 2012; Ireson et al., 2001). Due to the varying research results in regard to homogeneous placement and its overall academic effect, further research on this topic is required.

The results of this study indicated a significant difference among students who received instruction in homogeneously grouped classrooms and those who received instruction in heterogeneously group classrooms in the subject area of mathematics in regard to the Tennessee Comprehensive Assessment Program (TCAP) scores within the proficiency level in grades three, four, and five. However, the results of this study indicated no significant difference among students who received instruction in homogeneously grouped classrooms and those who received instruction in heterogeneously grouped classrooms in the subject area of English language arts. In this research study, two of the four null hypotheses were rejected.

Due to the misinterpretation of the definition of homogeneous grouping, Betts (2011) indicated that several disparities have occurred in respect to this research topic in the past. In comparison to the findings of Slavin (1990), in his meta-analysis of literature in regard to

homogeneous grouping, the findings of this research results were similar. When combining the research results of his study, Slavin determined there was not a connection between homogeneous grouping and the academic achievement of students in most subject areas.

Reuman (1989) and Wouters et al. (2012) noted that students do not score higher when homogeneously grouped. Even though some research has revealed positive academic benefits in respect to homogeneous grouping (Duflo et al., 2009), several research studies have noted areas such as class size, gifted programs, and schools identified as Title I as having a positive or negative impact. Some studies have indicated that students who performed above grade level showed an academic improvement when placed in gifted classes. Despite this fact, several gifted classes are only required for a small part of the day and students spend other parts of the day in heterogeneous grouped settings (Castambis & Buttaro, 2012). In order to allow students a chance to work with others with various capabilities, the utilization of heterogeneous grouped classes has become more popular in education.

The results of this research study were reported according to the type of ability grouping utilized, as well as the Tennessee Comprehensive Achievement Program (TCAP) scores. In contrast to the United States, the country of Finland does not rely on standardized testing to track the academic development of their students. Finland has focused its educational tenets on the following (Sahlberg, 2007):

- The majority of teachers in Finland have a master's degree or higher.
- The schools in Finland have a "culture of trust" between students, parents, teachers, and administrators.
- The leadership in Finland is supportive of students and teachers.

The tenets have led the educational system of Finland to not use standardized tests. The educators in Finland have been given more flexibility when creating lesson plans and curriculum. The added flexibility enables the educators more time to allow students to focus on reasoning and problem solving skills. Despite the fact that standardized testing has provided reliable data, it has not always indicated everything that the student has learned (Sackett et al., 2008).

As noted by Forgasz (2010), homogeneous grouping is implemented to help educators teach a class of students at a comparable instructional level. Hornby, Witte, & Mitchell, 2011 opined that some of the reasons that homogeneous grouping has been utilized in the educational setting were to meet the demands of standardized testing, meet the specific academic needs of students, and reduce the boredom of advanced students. While the above reasons are targeting the academic needs of students, research has indicated that, more often than not, grouping students into classrooms based on ability has not shown a significant difference, versus those who are in a heterogeneously grouped classroom.

Recommendations for Practice

Based on the findings of this study, the researcher provided the following recommendations for educational practice:

- When making instructional decisions, school administrators should consider the academic and societal effect on students. It is imperative that the whole child be considered, rather than just his or her academic growth.
- Teachers and administrators should meet to discuss the implications of grouping students by ability and how that would affect their students, based on the individual school demographics.

- When making instructional decisions, school administrators should consider the societal effect that homogeneous grouping will have on students before applying the practice. The social and academic benefits, as well as risks, must be taken into consideration when making instructional decisions.
- Due to the considerable effect that educators have on the academic achievement of students, educators should be a part of the decision-making process in respect to homogeneous grouping. If teachers agree in regard to instructional strategies, they will have a larger impact on student learning. Therefore, teachers have the ability to have a large impact on student learning and must be a part of the instructional strategies that are implemented.

Implications for Further Research

The results of this study have persuaded the researcher to make the following recommendations for further research in respect to student academic achievement and the type of grouping applied in the classroom:

- The research should be expanded by studying a larger number of schools in various parts of the United States to ascertain if the results are comparable to this study.
- The research should focus on student gains in academic performance for the four proficiency levels concerning the kind of grouping selected for instruction.
- Further research should focus on the variation of student grades and student academic achievement test scores between students who were homogeneously grouped and those who were not.

- Additional research should focus on various educational methodologies to determine the societal effects that homogenous grouping has on the self-efficacy of students and its impact on one's academic achievement.
- The research should focus on the comparison of the academic achievement of students, homogeneous grouping, and the overall effectiveness of the educator.
- Qualitative research methodologies may be utilized to determine the perception of educators in regard to homogeneous grouping and how it affects the teaching strategies of educators.
- Research should be conducted in respect to homogeneous grouping and the overall academic achievement of students who are enrolled in Title I schools across the United States.
- Additional research should be undertaken to investigate the difference in student growth and achievement, between students participating in a specific intervention or gifted program, and students who receive intervention and acceleration within a self-contained classroom.
- Research should be conducted to determine how the student's learning and human development are impacted instead of the student's academic achievement.
- Research should be conducted in respect to homogeneous grouping and the overall academic achievement of students who are not enrolled in Title I schools across the United States.

Conclusion

As schools make instructional decisions, it is necessary that administrators take into consideration the various research results in regard to the homogeneous grouping of students by

their academic capabilities, as well as the societal impact (Ireson & Hallam, 2001). In respect to this research study, the results have aligned with many other research studies which have indicated that there is not a significant difference in the standardized test scores of students who have been homogeneously grouped. This research study collected data from elementary schools from a rural county in East Tennessee. The elementary school administrators were asked to identify the type of grouping that was implemented in their schools for the 2014-2015 school year: homogeneously grouped or heterogeneously grouped. According to this research study, there was a statistically significant difference in grades three, four, and five in mathematics in the levels of below proficient and proficient between schools that utilized homogeneously grouped classrooms and those that were heterogeneously grouped by academic capability but not in English language arts. The results of this study were consistent with many others in finding that there were no significant differences among 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

- Ansalone, G., & Biafora, F. (2004). Elementary teachers' perceptions and attitudes to the educational structure of tracking. *Education, 125*(2), 249-258.
- Barker Lunn, J. (1970). *Streaming in the Primary School*. New York, NY: King, Thorne, and Stance Publishers.
- Betts, J. (2011) The Economics of tracking in education. *Handbook of the Economics of Education, 3*, 341-381.
- Boaler, J., William, D., & Brown, M. (2001). Students' experiences of ability grouping: Disaffection, polarization and the construction of failure. *British Educational Research Journal, 27*(5), 631-648.
- Bohrnstedt, G. (2013). Three decades of education reform: Are we still a nation at risk? American Institute for Research. Retrieved from www.air.org
- Burris, C., Heubert, J., Levin, H. (2006). Accelerating mathematics achievement using heterogeneous grouping. *American Educational Research Journal, 43*(1), 137-154.
- Byrne, B. (1988). Adolescent self-concept, ability grouping, and social comparisons: Reexamining academic track differences in high school. *Youth and Society, 20*(1), 46-67.
- Campbell, T. (2014). Stratified at seven: In-class ability grouping and the relative age effect. *British Educational Research Journal, 10*(5), 201-210.
- 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.eetsu.edu:3443/login?url=http://search.proquest.com/docview/131244223?accountid=10771>

- Catsambis, S., Mulkey, L., Buttaro, A., Steelman, L., & Koch, P. (2011). 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.
- Chorzempa, B., & Graham, S. (2006). Primary-grade teachers' use of within-class ability grouping in reading. *Journal of Educational Psych (ASCD)*, *98*(3), 529-541.
- Colangelo, N. & Davis, G. (1997). Handbook of gifted education (2nd ed.). *Gifted student quarterly*, *42*(3), 182-183.
- Collins, C. & Gan, L. (2013). Does sorting students improve scores? An analysis of class composition. Retrieved from <http://www.nber.org/papers/w18848>
- Csikszentmihalyi, M. & Wong, M. (2014). Motivation and academic achievement: The effects of personality traits and the quality of experience. *Journal of Personality*, *59*(3), 539–574.
- Davis, O. & Tracy, N. (1963). Arithmetic achievement and instructional grouping. *Arithmetic Teacher*, *10*(1), 12-17.
- Davies, J., Hallam, S., & Ireson, J. (2003). Ability grouping in the primary school: Issues arising from practice. *Research Papers in Education*, *18*(1), 45-60.
- Davidson, H. (2009). Ability Grouping. Retrieved from <http://www.education.com/reference/article/ability-grouping>
- Donaldson, M., LeChasseur, K. & Mayer, A. (2016). Tracking instructional quality across secondary mathematics and English Language Arts classes. Retrieved from <http://link.springer.com/article/10.1007/s10833-015-9269-x>

- Duflo, E., Dupas, P., & Kremer, M. (2009). Can tracking improve learning? *Education Next*, 9(3), 64-70.
- Elmore, R. (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 http://www.academia.edu/3499887/Corporate_Control_of_Public_School_Goals_High-Stakes_Testing_in_its_Historical_Perspective
- Emily, F., Robert, E., & Michael, K. (2003): The effect of ability grouping on student's achievements in science laboratory work. *Roeper Review*, 25(4), 212-220.
- 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
- 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>
- 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>
- 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.

- Gallagher, S. (2011). Teachers' perception of the socio-emotional development of intellectually gifted primary aged students and their attitudes towards ability grouping. *The Journal of the World Council for Gifted and Talented Children*, 26(2), 11-24.
- George, P. & Alexander, W. (2003). The exemplary middle school. *The Middle School Journal*, 20(1), 21-28.
- Green, S. & Salkind, N. (2011). Using SPSS for Windows and Macintosh: Analyzing and understanding data. Upper Saddle River, NJ: Prentice Hall
- Guisbond, L. & Neill, M., (2004). Failing our children: No child left behind undermines quality and equity in education. *Clearing House*, 78(1), 12-16.
- 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.
- Hopkins, G. (2008). Is ability grouping the way to go or should it go away? Retrieved from www.educationworld.com/a_admin/admin/admin009.shtml
- Hornby, G., Witte, C., & Mitchell, D. (2011), Policies and practices of ability grouping in New Zealand intermediate schools. *Support for Learning*, 26(3), 92-96.
- Ireson, J., & Hallam, S. (1999). Raising standards: Is ability grouping the answer? *Oxford Review of Education*, 25(3), 343-358.
- Ireson, J., & Hallam, S. (2001). Ability grouping in education. Thousand Oaks, CA. Sage Publications.

- Ireson, J., Hallam, S. & Plewis, I. (2001). Ability grouping in secondary schools: Effects on pupils' self-concepts. *British Journal of Educational Psychology*, 71(2), 315-326.
- Ireson, J., & Hallam, S. (2003). Secondary school teachers' attitudes towards and beliefs about ability grouping. *British Journal of Educational Psychology*, 73(3), 343–356.
- Kintz, M. (2011). Ability grouping and how it is affecting American classrooms. *ESSAI*, 9(20), 56-57.
- Klein, A. (2015). No child left behind: An overview. Retrieved from <http://www.edweek.org/ew/section/multimedia/no-child-left-behind>
- Kulik, J. (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., & Kulik, C. (1987). Effects of ability grouping on student achievement. *Equity and Excellence*, 23(1-2), 22–30.
- Kususanto, P. (2010). Students' self-esteem and their perception of teacher behavior: A study of between-class ability grouping. Retrieved from www.investigacion-psicopedagogica.org
- Lee, J. (2008). Is test-driven external accountability effective? Synthesizing the evidence from cross-state causal-comparative and co-relational studies. *Review of Educational Research*, 78(3), 608-644. Retrieved from <http://gse.buffalo.edu/gsefiles/images/Lee-RER.pdf>

- 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., Spence, J., Poulsen, C., Chambers, B., & Apollonia, S. (1996). With-in class grouping. A meta-analysis. *Review of Educational Research*, 66(3), 423-458.
- Loveless, T. (2009). Tracking, detracking: High achievers in Massachusetts middle school Washington, DC: Thomas B. Fordham Institute.
- Loveless, T. (2013). The resurgence of ability grouping and persistence of tracking. Retrieved from <http://www.brookings.edu/research/reports/2013/03/18-tracking-ability-grouping-loveless>
- Matthews, M., Ritchotte, J. & McBee, M. (2013). Effects of school wide cluster grouping and within-class ability grouping on elementary school students' academic achievement growth. *High Ability Studies*, 24(2), 81-97.
- Marzano, R., Welch, L., Adams, G., Brown, J., & Welch, A. (2008). Association for Supervision and Curriculum Development. The Art and science of teaching. Alexandria, VA: ASCD.
- McCarter, A. (2014). Ability grouping in elementary education (Doctoral dissertation). Retrieved from www.dc.etsu.edu/etd/2459
- McGuinn, P. (2011). Educational policy stimulating reform: Race to the top, competitive grants and the Obama education agenda. *Educational Policy*, 26(1), 136-159.
- McMillan, J., & Schumacher, S. (2010). Research in education: Evidence-based inquiry. Boston, MA: Pearson.

- Neihart, M. (2007). The socio-affective impact of acceleration and ability grouping. *Gifted Child Quarterly*, 51(4), 330-341.
- Nichols, S. & Berliner, D. (2007). *Collateral damage: How high-stakes testing corrupts America's schools*. Cambridge, MA: Harvard Education Press.
- No Child Left Behind Act (NCLB) Act of 2001, Pub. No.107-110, 20 U.S.C. (2002).
- Pekrun, R., Goetz, T., Titz, W. & Perry, R. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, 37(2), 91-106.
- Pierce, R., Cassady, J., & Adams, C. (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.
- Plunkett, M., & Kronberg, L. (2011). Learning to be a teacher of the gifted: The importance of examining opinions and challenging misconceptions. *Gifted and Talented International*, 26(1): 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.
- Preckel, F. & Brull, M. (2008). Grouping the gifted and talented: Are gifted girls most likely to suffer the consequences. *Journal for the Education of the Gifted*, 32(1), 54-85.

- Preckel, F., Gotz, T., and Frenzel, A. (2010). Ability grouping of gifted students: Effects on academic self-concept and boredom. *British Journal of Educational Psychology*, 80(3), 451–472.
- Reuman, D. (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.
- Rist, R. (1970). Student's social class and teacher expectations: The self-fulfilling prophecy in ghetto education. *Harvard Educational Review*, 40(3), 411-451.
- 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*, 30(2), 141-180.
- Rubin, B. (2008). De-tracking in context: How local constructions of ability complicate reform. *The Teachers College Record*, 110(3), 9-10.
- Sackett, P., Borneman, M., & Connelly, B. (2008). High stakes testing in higher education and employment: appraising the evidence for validity and fairness. *American Psychology*, 63(4), 215.
- Sahlberg, P. (2007). Education policies for raising student learning: The Finnish Sattler approach. *Journal of Education Policy*, 22(2), 173-197.
- Sanders, W., Wright, S., & Horn, S. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, 11(1), 57-67.

- Sattler, J. (2001). Issues related to the measurement and change of intelligence. In assessment of children: Cognitive applications (4th ed.). San Diego, CA: Jerome M. Sattler, Publishing Inc.
- Shield, C. (2010). A comparison study of student attitudes and perceptions in homogeneous and heterogeneous classrooms. *Roeper Review*, 24(3), 115-119.
- Slavin, R. (1986). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. Report No. 1. Baltimore, Md.: Johns Hopkins University, Center for Research on Elementary and Middle Schools.
- Slavin, R. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research*, 57(30), 293-336.
- Slavin, R. (1990). Achievement effects of ability grouping in secondary schools: A best evidence synthesis. *Review of Educational Research*, 60(3), 471-499.
- Snider, V. & Schumitsch, R. (2006). A Comparison of teacher attitudes and beliefs about issues in education across conventional and direct instruction schools. *Journal of Direct Instruction*, 6(1), 17-33.
- Tach, L., & Farkas, G. (2006). Learning-related behaviors, cognitive skills, and ability grouping when schooling begins. *Social Science Research*, 35(4), 1048-1080.
- Tieso, C. (2003). Ability grouping is not just tracking any more. *Roeper Review*, 26(1), 29-37.
- Totten, I., & Bosco, K. (2008). The impact of ability grouping in introductory geology labs. *Geological Society of America Abstracts with Programs*, 40(3), 163.

- Trigg-Smith, R. (2011). Investigating teachers' impact on grouping practices: Justifying an analysis on how teachers incorporate ideas of equity into grouping practices. *Educate Journal*, 11(1), 28-44.
- Vialle, W., Heaven, P. & Ciarrochi, J. The relationship between self-esteem and academic achievement in high ability students: Evidence from the Wollongong Youth Study. *Australasian Journal of Gifted Education*, 24(2), 17-23.
- Venkatakrishnan, H. and William, D. (2013). Tracking and mixed-ability grouping in secondary school mathematics classrooms: A case study. *British Educational Research Journal*, 29(2), 189–204.
- Vogl, K. & Preckel, F. (2014). Full-time ability grouping of gifted students impacts on social self-concept and school-related attitudes. *Gifted Child Quarterly*, 58(1), 51-68.
- Willard, C. (1996). Liberalism and the problem of knowledge. A new rhetoric for modern democracy. Chicago, IL. University of Chicago Press
- 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.

Appendices

Appendix A: Request to Conduct Research Study

Appendix A

Request to Conduct Research Study

Hello Mr. X,

I am currently working on my doctorate at Carson Newman. I am writing the dissertation concerning ability grouping in elementary schools with a focus on 3rd,4th, and 5th grades. For the research study, I plan to compare Tennessee Comprehensive Assessment Program (TCAP) scores from the 2014-2015 school year of ability grouped classes to non-ability grouped classes.

May I have your permission to contact the elementary school principals of X County as part of my research, please? I plan to send the principals questions concerning whether their 3rd,4th, and 5th grade classes are ability grouped or not. When I begin to compile the data, I will go to the Tennessee Department of Education website to compare test scores.

I appreciate your time and attention concerning this matter. I look forward to hearing from you soon.

Thanks so much!
Angela Johnson, Ed.S.

Appendix B: Approval Response from the Director of Schools in X County.

Appendix B

Approval Response from the Director of Schools in X County

Angela,

Yes, I will approve for you to send a questionnaire to the elementary principals regarding your research. A principal's participation shall be voluntary. Best of luck with your study and completion of your degree. Thanks.

Mr. X
Director of Schools
X County Schools
200 N. Depot St.
Rogersville, TN 37857
423-272-7629 ext. 2023
Fax – 423-272-2207

Appendix C: Initial Email to Participants

Appendix C

Initial Email to Participants

Hello everyone! How are you?

I am currently working on my doctorate at Carson Newman University. I am focusing my research on ability or homogeneous grouping in elementary schools with an emphasis on 3rd,4th, and 5th grades. For the research study, I plan to compare Tennessee Comprehensive Assessment Program (TCAP) assessment scores from the 2014-2015 school year of ability grouped classes to non-ability grouped classes. When I compile the data, I will go to the Tennessee Department of Education website to compare test scores.

Would you please take the time to answer the optional six questions at the end of this email?

Thanks so much!
Angela Johnson

Appendix D: Questions for Participants

Appendix D

Questions for Participants

1. Do the 3rd grade class or classes of your school group by ability in math?

Yes No Not Applicable

2. Do the 3rd grade class or classes of your school group by ability in reading?

Yes No Not Applicable

3. Do the 4th grade class or classes of your school group by ability in math?

Yes No Not Applicable

4. Do the 4th grade class or classes of your school group by ability in reading?

Yes No Not Applicable

5. Do the 5th grade class or classes of your school group by ability in math?

Yes No Not Applicable

6. Do the 5th grade class or classes of your school group by ability in reading?

Yes No Not Applicable