

EFFECTS OF GROUPING ON 4TH GRADE
MATHEMATICS ACHIEVEMENT

by

Brian Marinelli

A dissertation

submitted in partial fulfillment

of the requirements for the degree of

Doctor of Education in Curriculum and Instruction

Boise State University

December 2017

© 2017

Brian Marinelli

ALL RIGHTS RESERVED

Brian Marinelli

SOME RIGHTS RESERVED



This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 United States license. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/us/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the dissertation submitted by

Brian Marinelli

Dissertation Title: Effects of Grouping on 4th Grade Mathematics Achievement

Date of Final Oral Examination: 12 October 2017

The following individuals read and discussed the dissertation submitted by student Brian Marinelli, and they evaluated his presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Jonathan Brendefur, Ph.D.	Chair, Supervisory Committee
Carl F. Siebert, Ph.D.	Member, Supervisory Committee
Sara Hagenah, Ph.D.	Member, Supervisory Committee
Sherry Dismuke, Ed.D.	Member, Supervisory Committee

The final reading approval of the dissertation was granted by Jonathan Brendefur, Ph.D., Chair of the Supervisory Committee. The dissertation was approved by the Graduate College.

DEDICATION

This dissertation is dedicated to my loving and supportive family. Throughout my entire life I knew that my parents, Mike and Emma, were in my corner and proud of me every day. My loving daughter, Sara, has showered me with love and smiles every day of her life. My son Dante has made me a better teacher and father since the moment he came into my life. Finally, I want to thank the Lord for not only allowing me to achieve all I have but for bringing my wife Midori into my life. She has provided love and support every moment of this long journey.

ACKNOWLEDGEMENTS

I would like to express sincere gratitude to my entire committee. I would like to thank Dr. Sherry Dismuke for keeping me focused with her passionate advocacy for all students. I would also like to thank Dr. Sara Hagenah for not only helping me to the finish line with my dissertation, but pressing me to think bigger. I would not have been able to complete this milestone without Dr. Carl Siebert. His guidance and feedback was instrumental in helping me with my methodology. Dr. Jonathan Brendefur has been more than a committee chairman. His mentorship over the past couple of years has helped me to become a better mathematics teacher and a researcher. Finally, I would like to thank my wife, Midori. This milestone would not have been achieved without her love and encouragement.

ABSTRACT

Questions involving the equitable distribution of mathematics instruction have been addressed since at least the early 1990's. Since this time, little research has been conducted on the antecedents and effects of grouping elementary school students within homogenous mathematics groups. The purpose of this quantitative study was to compare the mathematics achievement of 4th grade elementary school students who were grouped in either homogenous or heterogeneous mathematics classes. A causal-comparative design was utilized in an attempt to find relationships between the independent variables of mathematics grouping level, sex, and ethnicity and the dependent variable of mathematics achievement. Data were analyzed using independent-sample t-tests, analysis of variance (ANOVA) and a multiple linear regression. Grouping was found to have little statistical significance on the mathematical achievement of 4th grade students within this study. The independent variables of sex and ethnicity were significantly associated with the dependent variable of mathematical achievement. There are many opportunities to further study the antecedents and effects of grouping elementary school students within homogenous mathematics groups. Other areas of research are: 1) what factors are used to group students into mathematics ability groups within elementary schools, 2) what factors determines how teachers are assigned to teach different levels of mathematics within elementary schools, and 3) at what grade level does sorting begin in mathematics?

Keywords: mathematics equity, grouping, sorting, tracking

TABLE OF CONTENTS

DEDICATION	v
ACKNOWLEDGEMENTS	vi
ABSTRACT	vii
LIST OF TABLES	xiii
LIST OF FIGURES	xv
CHAPTER 1: INTRODUCTION	1
Introduction	1
Background of the Study	3
Problem Statement	5
Purpose of the Study	6
Nature of the Study	6
Research Questions and Hypotheses	8
Definition of Terms	9
Assumptions	10
Limitations	10
Scope	11
Delimitations	11
Significance of the Study	11
Organization of the Study	12

CHAPTER 2: LITERATURE REVIEW	14
Introduction.....	14
Mathematics Equity	15
What is Mathematics Equity?.....	15
“Math is Neutral” Viewpoint.....	20
Role of Math as a Gatekeeper in Society.....	21
Changing Demographics in U.S.	23
Achievement vs. Opportunity Gap	24
Teacher Qualifications	28
Rich Contextual Problems	30
Access to Advanced Courses	31
Educational Debt.....	32
Social Justice.....	33
Critical Perspective	34
Using Math to Challenge Authority.....	35
Valorization of Mathematics.....	37
Funds of Knowledge.....	39
Communities of Practice.....	40
Tracking, Grouping, and Sorting	42
Teacher Judgement	45
Teachers Evaluations	45
Naturalistic Expectations	47
Guiding Theories	48

Lack of Current Research	52
What Needs to be Studied.....	52
Research Question	53
CHAPTER 3: METHODS AND PROCEDURES	55
Purpose.....	55
Research Design.....	56
Setting and Sample	57
Population Definition.....	57
Sampling Method.....	57
Participants.....	58
Instrumentation	60
Predictor/Independent Variables.....	60
Outcome/Dependent Variable.....	60
Data Collection and Analyses.....	61
Data Collection	61
Data Analysis	62
CHAPTER 4: RESULTS.....	69
Overview of the Study	69
Data Analyses	70
Data Cleansing	70
Research Questions #1 - Grouping	72
Grouping	73
Research Questions #2 – Ethnicity and Sex	78

Ethnicity	78
Sex.....	82
Sex and Ethnicity	84
Sex, Ethnicity, and Grouping.....	85
CHAPTER 5: CONCLUSION	95
Summary	95
Purpose.....	95
Connection to Back to Literature Review.....	96
Methodology	97
Interpretation of Findings	99
Grouping	99
Ethnicity	101
Sex.....	102
Ethnicity, Sex, and Grouping.....	103
Practical Implication for Educators	103
Recommendations for Further Study	105
Conclusion	105
REFERENCES	107
APPENDIX A.....	118
4th Grade Mathematics Grouping Online Teacher Questionnaire	119
APPENDIX B	134
Independent Variables	135
APPENDIX C	137

Teacher Recruitment Script and Consent	138
APPENDIX D.....	140
Modified Teacher Recruitment Script and Consent.....	141

LIST OF TABLES

Table 1:	Belenky's (1986) Epistemological Powers of the Individual	51
Table 2:	Respondent Rate	61
Table 3:	Valid Assessments	63
Table 4:	Change_SS vs Ethnicity.....	64
Table 5:	Skewness & Kurtosis	64
Table 6:	Test of Normality - Sex.....	65
Table 7:	Test of Normality - Ethnicity	66
Table 8:	Test of Normality - Grouping	66
Table 9:	Test of Normality - Grouping Level	67
Table 10:	T-Test Post hoc Power Computation	70
Table 11:	Valid Assessments	71
Table 12:	Sex.....	71
Table 13:	Change_SS vs Ethnicity.....	72
Table 14:	Achievement vs. Grouping	74
Table 15:	Change_SS vs Levels of Grouping.....	78
Table 16:	Ethnicity vs Change_SS.....	80
Table 17:	Ethnicity and Grouping vs Change_SS.....	81
Table 18:	Ethnicity and Grouped vs Change_SS	82
Table 19:	Sex vs Change_SS	83

Table 20:	Sex and Grouping vs Change_SS	84
Table 21:	Sex and Ethnicity vs Change_SS	85
Table 22:	Sex, Ethnicity, and Grouped vs Change_SS.....	86
Table 23:	Sex, Ethnicity, and Grouping Level vs Change_SS	87
Table 24:	Regression Correlations	93
Table 25:	Residuals Statistics.....	93
Table 26:	Casewise Diagnostics.....	94

LIST OF FIGURES

Figure 1:	Scatterplot of Residuals	90
Figure 2:	Histogram of Residuals.....	91
Figure 3:	Plot of Regression Standardized Residuals.....	92

CHAPTER 1: INTRODUCTION

Introduction

Almost three decades ago, Stanic (1989) challenged all mathematics educators with the statement, “If mathematics educators take seriously the goal of equity, they must question not just the common view of school mathematics but also their own taken-for-granted assumptions about its nature and worth.” Conversations that revolve around what are the mathematics needs of this nation’s students should focus not only on the curriculum, but also how we can ensure all students have the opportunity to learn (Schoenfeld, 2002). The issue of equitable access to mathematics instruction has been addressed by the National Council of Mathematics Teachers’ equity principle that states that all students are entitled to highly qualified teachers and rich mathematics instruction (Bartell, 2007; Mathematics, 2000). Several studies found that teacher quality is inequitably distributed across all measures of student disadvantage (Goldhaber, Lavery, & Theobald, 2015; Hoogeveen, Van Hell, & Verhoeven, 2005).

Research into mathematics equity is a complex undertaking because it is situated within broader educational, social, and political contexts that include issues of race, class, sex, language, culture, and power (Bartell, 2007). The factors that aggregate to form cultural identity are country of origin, language, and/or religion. These factors interact in a much more complex way than due skin color or last name. In most instances, classroom

teachers are on the front line of this ever-changing climate of cultural diversity and pedagogy.

National and international mathematics assessments give insight to the widening achievement gap in the U.S. Darling-Hammond (2010) states of all the nations that took the 2006 PISA (Program in International Student Assessment), the US ranked 21 out of 30 nations in science and 25 out of 30 nations in mathematics. Asian and White students from the U.S. scored above the PISA average while African American and Hispanic students from the U.S. scored below the average. Overall, Whites perform much better on these assessments than do Hispanics who, in turn, do slightly better than African Americans. This is similar to the results of the 2012 PISA. The United States performed below average in mathematics in 2012 and is ranked 27th (PISA, 2012). Socio-economic disadvantage had a large impact on how the United States performed. 15% of the variation in student performance was explained by students' socio-economic status. This is in contrast with less than 10% in other similar countries. Even more alarming is the fact that in the U.S., as a student gets older, these gaps also keep increasing (Secada, 1992). The achievement gap between races and different ethnic groups has been the focus of educational researchers since the early-1990's. Secada's (1992) seminal work "Race, Ethnicity, Social Class, Languages, and Achievement in Mathematics" helps to shed light on the state of affairs of the "differentially effective" educational system within this country. Student's receive varying educational experiences based on social class, race, ethnicity, language background, sex, and other demographic characteristics (p. 623).

One possible structure attributed to the widening achievement gap is that of homogenous grouping and tracking of students. Increased time on mathematics and the

taking of advanced coursework are two of the most powerful predictors of student achievement (Secada, 1992). Over twenty-five years ago Oakes (1990) stated:

Some see schools as meritocratic institutions that consider achievement itself as the principal mediator of opportunity, arguing that children who achieve more are better able to benefit from and more deserving of the limited resources that are available. Others explain opportunity, achievement, and participation as a function of mental capacity; for them, the most important opportunities are conferred at birth or before. (p. iii)

Oakes statement is a sobering thought. If it is to be believed, then achievement begets achievement. Those who exhibited high achievement are given more educational opportunities.

Background of the Study

Grouping and tracking students inevitably separates students by race, ethnicity, native language, and class (Loveless & Diperna, 2000). In the past, Latino students who scored within the 60th percentile were 50 percent less likely to be placed into a college preparatory class when compared to White and Asian students (Oakes, 1992). Not only are non-Asian minority and poor students tracked into lower mathematics courses, but those courses tend to have less qualified teachers (Secada, 1992). This trend is still true today, with Black, Hispanic and poor children still frequently tracked into remedial classes, while middle-class White children are tracked into honors courses (Loveless & Diperna, 2000).

Tracking is put into place because educators believe and argue that it better meets the individual needs of students. The antecedent of tracking students into low-ability mathematics classes at the secondary level typically begins with recommendations initiated by the student's upper elementary school teacher. Additionally, there is the belief that students are better served later in life by putting students into homogenous

mathematics groups. Additionally, there is widespread belief that children's intellectual difference is so great and that there is such a perceived varying ability levels that they need to be taught in separate classrooms (Oakes, 1987 & 1990).

Typically, students are grouped based on judgements of their ability and educability (Secada, 1992). Secada (1992) goes on to state that there is evidence that schools do not sort students only on merit and they do not sort students uniformly. There are substantial ethnic differences in how students are assigned to tracks, even after adjusting for prior achievement.

Hattie's meta-analyses (2008) has found that tracking not only has minimal effects on learning outcomes but has profound negative equity effects. The resistance to elimination of tracking programs is typically from "high-track" teachers and parents who have benefited from tracking (Mathis, 2013). The goal of curricular stratification (tracking) in American schools has been to create homogeneity based on student ability at the classroom level (Burris, Welner, and Bezoza, 2009). Research has shown that tracked classes have a wide range of student ability levels due to how enrollment decisions are made. These decisions include not only test scores and prior achievement, but student behavior, parent preference, completion of prerequisites, teacher judgment, and counselor guidance (Useem, 1992). Not only is the selection process for tracking students suspect, but teacher qualification in low-track classes is a concern. Low-track courses frequently had more out-of-field teachers than do high-track courses (Tate, 2008). Stinson (2004) gives a qualitative insight into what tracking looks like at the classroom level.

My experiences as a secondary mathematics teacher, preservice-teacher supervisor, and researcher supported Oakes's (1985) assertions that often students are distributed into "ability" groups based on their race, sex, and class. Nonetheless, my perception after five observations was that ability grouping

according to these attributes was diminishing—at least in these elementary schools. In other words, the student make-up of each mathematics lesson that I observed appeared to be representative of the demographics of the school.

However, on my sixth observation, at an elementary school with 35.8 % Black, 12.8 % Asian, 5.3 % Hispanic, 3.5 % Multi-racial, and 0.5 % American Indian children, I observed a 3rd grade mathematics lesson that was 94.4% White (at least it was 50% female). The make-up of the classroom was not initially unrepresentative of the school's racial/ethnic demographics, but became so shortly before the start of the mathematics lesson as some students left the classroom while others entered. When I questioned why the students were exchanged between classrooms, I was informed that the mathematics lesson was for the “advanced” third graders. (p.8)

Problem Statement

A key component of grouping students into differentiated mathematics classes is a teacher's ability to judge the readiness level of a student. Teacher's perception of a class's ability plays an important role in deciding what to teach and how to teach it. Judgments about academic ability often lead to the segregation of students into separate elementary and middle-school classes and to enrollment in different senior high school courses (Oakes, 1990, p. 17). Compounding the detrimental effects of tracking is that teachers are unable to accurately determine which students are qualified for or would benefit from more advanced mathematics courses. Factors such as SES, parent influence, and racial/cultural membership are frequently used as sorting mechanisms (Darling-Hammond, 2010; Oakes, 1990; Secada, 1992). Teacher judgments are frequently used for instructional planning, screening, placement, and referrals (Martinez, Stecher, & Borko, 2009).

There is currently a lack of research into the structural processes that upper elementary school teachers use to recommend students for advanced mathematics courses and what the effects of this grouping has on students' mathematical achievement. Students moving into the middle school grades (6 through 8) are often placed within

homogeneous mathematics groups based on several factors, one of which is recommendations made by their previous elementary school mathematics teacher. The factors that elementary teachers use to make these recommendations may vary from perceptions of students' work ethic, demographics (i.e. sex, race, cultural identity, SES, etc...), mathematics ability, and student home life.

Purpose of the Study

The purpose of this quantitative study was to compare the mathematics achievement of 4th grade elementary school students who were grouped in either homogenous or heterogeneous mathematics classes. Student achievement was measured from the beginning of 4th grade till the 4th quarter of the school year based on the STAR standardized mathematics assessment. Variations in growth were control for by student ethnicity, student sex and with the primary interest on the effects of ability grouping of students on mathematics achievement.

Nature of the Study

This study uses a causal-comparative design in an attempt to find relationships between the independent and dependent variables. A causal-comparative design's goal is to determine whether the independent variable affected the outcome variable by comparing two or more groups of individuals after the event has occurred (*ex post facto*) (Salkind, 2010).

The causal-comparative design lends itself to studies in which independent variables are unable to be manipulated, as is frequently the case with educational research. Variables such as sex, ethnicity, and socioeconomic levels are already set and cannot be decided. Another key aspect to causal-comparative designs is subjects are

already in groups in contrast to a true experimental research design where subjects are randomly selected (Salkind, 2010).

According to Salkind (2010), inferential statistical methods, such as chi-square test, paired-samples and independent t tests, and analysis of variance (ANOVA), are appropriate within a casual-comparative research study when the “researcher hopes to demonstrate that a relationship exist between the independent and dependent variables.”

Crosstabs and chi-square test are used to compare two nominal variables, a nominal variable with an ordinal variable, or two ordinal variables (Muijs, 2010). The distribution of females and males across the student categories of general education, special education, and limited English proficiency were explored using crosstabs. The distribution of ethnicity across student categories was also evaluated. The statistical significance of these distributions was analyzed using chi-square test.

T-tests are appropriate when comparing the differences between the means of a continuous variable between two groups (Muijs, 2010). Two different continuous dependent variables were analyzed. The first is the change in the student’s scaled score (change_SS) that was measured from their fall STAR mathematics assessment and that of their spring score. The second dependent variable is change in a student’s grade equivalency (change_GE) that was measured during the same time frame as their scaled score. Independent sample t-tests were conducted to check the statistical significance and effect size between the dependent variable (mathematics achievement) and the independent variables of sex and whether or not students were grouped in mathematics.

Analysis of variance (ANOVA) allows comparison of a continuous dependent variable and several groups (Muijs, 2010). An F-test can be used *post hoc* to determine

not only the statistical significance between groups, but the effect size. Main effects and interactions were analyzed on the scale score change (change_SS) dependent variable and the independent variables of sex and ethnicity.

A multiple linear regression was conducted to assist in measuring the association between the independent variables of sex, ethnicity, and grouping level and the dependent variable, change in scale score (change_SS). This will help in the construction of a linear equation that will predict values of the dependent variable based on the independent variables (Gray & Kinnear, 2012).

Research Questions and Hypotheses

There were two guiding research questions and associated hypotheses for this study.

Research Question 1 - Does ability grouping affect mathematics achievement of 4th grade students?

H₀: Ability grouping does not have an affect on mathematics achievement of 4th grade students.

H₁: Ability grouping does have an affect on mathematics achievement of 4th grade students.

Research Question 2 - Does ability grouping affect the mathematics achievement of 4th grade students differently based on sex and ethnicity of students?

H₀: Ability grouping by mathematics achievement of 4th grade students does not affect students differently by sex and ethnicity.

H₁: Ability grouping by mathematics achievement of 4th grade students affects students differently by sex and ethnicity.

Definition of Terms

The operational definitions for this study were as follows:

Race/ethnicity: Indicates general racial or ethnic heritage. Based on Hispanic ethnic category and five single-race categories found below. (Musu-Gillette et al., 2016).

White: Students having origins in any of the original peoples of Europe, the Middle East, or North Africa.

Black or African American: Students having origins in any of the black racial groups of Africa.

Hispanic or Latino: Students of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin.

Asian: Students having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian Subcontinent.

Native Hawaiian or Other Pacific Islander: Students having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.

American Indian or Alaska Native: Students having origins in any of the original peoples of North, South, and Central America, and who maintains tribal affiliation or community attachment.

Grade Equivalent (GE): Measure of what a normal grade placement is of a student of which a particular score is typical. Individual months are in tenths, ranging from 0.0 to 12.9+. (Learning, 2015).

Limited English Proficiency (L.E.P.) (NCES, 2016): Students that are ages 3

through 21, who were not born in the United States or whose native languages are other than English, and whose difficulties in speaking reading, writing or understanding the English language may be sufficient to deny the individuals the ability to be successful in a classroom where English is the instructional language (NCES, 2016).

Scaled Score (SS): Test scores are converted into scale scores by first “estimating each student’s location on the Rasch ability scale, based on the difficulty of the item and by the pattern of right and wrong answers. It then uses a linear transformation to make all scores positive integers ranging from 0 to 1400 (Learning, 2015).”

Special Education (S.P.E.D.): Students who receive specifically designed instruction that meet the unique needs of those with a disability, including classroom instruction, instruction in physical education, home instruction, and instruction in hospitals and institutions (Packer, 2002).

Assumptions

It is assumed that all participants answered the online survey honestly and to the best of their abilities.

The assumptions associated with the statistical analysis of this study are covered in Chapter 3: Methods and Procedures.

Limitations

The scope of this study was limited by both resources and time. Dependent or confounding variables such as school size, student socioeconomic status, motivation, and effort were not addressed within this study. Additionally, teacher bias towards type of group being taught was not controlled for. This study did not account for 4th grade students that moved between ability groups during the school, were enrolled in more than

one math class at a time, or receive after-school mathematics assistance. Due to the small sample available for the study, results may not be generalizable beyond the specific population from which the sample was drawn.

Scope

In this study, a causal-comparative research design was used to explore the effects that homogenous mathematics grouping had on the achievement of 4th grade students. Students mathematics scores on a nationally normed assessment were compared from the beginning of the school year to the spring of the same school year.

Delimitations

The sample population for this study focused on one school district due to time and cost constraints. This school district was willing to provide access to all of its 4th grade teachers and elementary schools. An additional delimitation was the availability for a nationally normed pre- and post- assessment of mathematics achievement. The sample school district had already been conducting assessment of its 4th grade students using the STAR mathematics assessment.

Significance of the Study

This study provides valuable insight into the limited benefits gained from homogenously grouping 4th grade within mathematics classrooms. The results and discussion adds to the limited body of literature that concerned with the ability grouping of mathematics students within the elementary grades. Educational researchers with a focus on mathematics equity, in particular, those concerned with decreasing the opportunity gap of female and minority students, can benefit from the insights of this study.

Classroom teachers and school administrators spend valuable time reviewing class rosters at the beginning of each school year. Frequently the discussion turns to how to structure the students' day so then can be grouped into homogenous ability groups in mathematics. This study is a jumping off point for those discussions.

Organization of the Study

Chapter one provides an introduction and significance to this study. The primary focus is to explore the effects that ability grouping mathematics students, specifically 4th grade students. Additionally, the interaction between grouping, sex, and ethnicity has on mathematics achievement will be studied.

Chapter two is comprised of a comprehensive review of literature of mathematics equity that spans several decades. Literature was looked at through three theoretical lenses. The first lens is how mathematics equity is defined, its role in society, and the influence on both the achievement and opportunity gap. The term equity is value-laden and frequently used by authors to mean different things. It is important to begin by first delineating how "equity" is used within academic scholarship. The second theoretical lens was through a social justice viewpoint. Mathematics equity is approached as a means to social justice. The final theoretical lens was through a teacher judgement lens.

Chapter three will outline the methodology of this causal-comparative design. This quantitative design's goal is to determine whether the independent variable affected the outcome variable by comparing two or more groups of individuals after the event has occurred (ex post facto) (Salkind, 2010). In this case, mathematics grouping, sex and ethnicity will be analyzed to determine their effects, if any, on the mathematics achievement of 4th grade students.

Chapter four is an analysis of the data collected in the course of this study. According to Salkind (2010), inferential statistical methods, such as chi-square test, paired-samples and independent t tests, and analysis of variance (ANOVA), are appropriate within a casual-comparative research study when the “researcher hopes to demonstrate that a relationship exist between the independent and dependent variables.” In addition to these inferential methods, a linear regression was also conducted.

Chapter five provides conclusions, discussion, and recommendations. This chapter synthesizes the analysis from chapter four. The goal of this chapter to provide substantial insight into the effects of grouping 4th grade mathematics students for the benefit of researchers, classroom teachers, and school administrators.

CHAPTER 2: LITERATURE REVIEW

Introduction

The literature pertaining to research in the field of mathematics education extends back at least fifty years. Questions involving issues of culturally diversity and the equitable distribution of mathematics education received heightened awareness in the 1990's and 2000's. Researchers like Darling-Hammond (2010), Gutiérrez (2007), Ladson-Billings (1995, 1997, 2006), Lubienski (2002, 2008), Moses and Cobb (2001), Oakes (1985, 1987, 1990, 1992), Secada (1989, 1992, 1995) and Secada, Fennema, and Adajian (1995) have attempted to strike a pragmatic chord from the critical notes of Apple (2004), Bourdieu (1977), Freire (2000), and Žižek (2009). The purpose of this literature review is to provide an overview of not only the antecedents of grouping elementary school students in mathematics, but what the effects of grouping are to mathematics achievement.

Literature was looked at through three theoretical lenses. The first lens is how mathematics equity is defined, its role in society, and the influence on both the achievement and opportunity gap. The term equity is value-laden and frequently used by authors to mean different things. It is important to begin by first delineating how "equity" is used within academic scholarship. The second theoretical lens was through a social justice viewpoint. Mathematics equity can be approached as a means to social justice that allows students access to the tools needed to "thrive in the 21st century (Bond & Chernoff, 2015). The final theoretical lens was through a teacher judgement lens. The

goal of this literature review was to look at mathematical equity as an issue of social justice through which practices at the school and teacher level be addressed.

Mathematics Equity

Mathematics equity is a complex social construct that will be explore within this section. The term equity is frequently described by educators, policy makers, and researchers to indicate a “fairness” within classroom (Reed & Oppong, 2005, p. 4). Some relate equity with a lack of quality, a “dumbing down” of the curriculum to meet the needs of the struggling student (Benbow & Stanley, 1996). The vast amount of literature written by educational researchers over the last 30 years have tried to make a case that mathematics equity and quality can coexist.

This section will explore various operational definitions of mathematics equity, the view that mathematics has no hegemonic value (neutral subject), the role of mathematics in society as a “gatekeeper” to future academic and economic success, the changing demographics of the United States, and the difference between the achievement gap and the opportunity gap.

What is Mathematics Equity?

Almost three decades ago, Stanic (1989) challenged all mathematics educators with the statement, “If mathematics educators take seriously the goal of equity, they must question not just the common view of school mathematics but also their own taken-for-granted assumptions about its nature and worth.” Conversations that revolve around what are the mathematics needs of this nation’s students should focus not only on the curriculum, but also how we can ensure all students have the opportunity to learn (Schoenfeld, 2002). The issue of equitable access to mathematics instruction has been

addressed by the National Council of Mathematics Teachers' equity principle that states that all students are entitled to highly qualified teachers and rich mathematics instruction (Bartell, 2007; Mathematics, 2000). Several studies found that teacher quality is inequitably distributed across all measures of student disadvantage (Goldhaber et al., 2015; Hoogeveen et al., 2005).

Confounding the research of mathematics equity is the ambiguous nature of cultural and race. Ethnic and racial categories are frequently self-assigned (Secada, 1992). Labels such as Negro, Black, Afro-American, and African American can be thought of as interchangeable, except to those who identify within those groups. The same can be said of Hispanic and Asian groups. In order to provide clarification and standardization for researchers, the National Center for Education Statistics publishes definitions of both ethnicities and race.

The ethnicity of Hispanic or Latino is based on a person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. The term "Spanish origin" can be used in addition to "Hispanic or Latino." Race is based on five different categorizations. American Indian or Alaska Native is a person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment. The Asian race is based on a person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent, including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam. Black or African American is based on a person of origin from any of the Black racial groups of Africa. Native Hawaiian or Other Pacific Islander is a person with origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands. Finally, someone considered of the White race is based on the origins in any of the original peoples of Europe, the Middle East, or North Africa. (National Center for Education Statistics, 2002)

The ability to accurately define both race and ethnicity is instrumental when exploring educational issues like mathematics equity.

Research into mathematics equity is a complex undertaking because it is situated within broader educational, social, and political contexts that include issues of race, class, sex, language, culture, and power (Bartell, 2007). The factors that aggregate to form cultural identity are country of origin, language, and/or religion. These factors interact in a much more complex way than due skin color or last name. In most instances, classroom teachers are on the front line of this ever-changing climate of cultural diversity and pedagogy.

The definition of “equity” needs to be viewed within complex cultural, social, and political spheres. Gutiérrez (2007) states of equity being more of a process of justice than of equality. She defines mathematics equity as: 1) being unable to predict students’ mathematics achievement and participation based solely on race, class, ethnicity, sex, beliefs, and language proficiency, 2) being unable to predict students’ ability to analyze, reason about, and especially critique knowledge and events in the world as a result of mathematics practice, based solely on race, class, ethnicity, sex, beliefs, and language proficiency, and finally 3) an erasure of inequities between people, mathematics, and the globe. The notion that educators should be unable to predicate mathematics achievement based on ethnicity or sex is the focus of this study.

Cobb and Hodge (2010) offer a different definition in which to view equity within a mathematics classroom. Their three-part definition of equity begins with Bruner’s (1986) assertion that mathematics reasoning enables students to have “clout” or to participate in significant out-of-school practices in relatively substantial ways. They further state that the notion of equity is that schooling should “compare and differentiate” between students in ways that enhance their future academic and economic opportunities.

They finally assert that mathematics equity should lead to the “cultivation of students’ interests in and feelings of efficacy about the future study of mathematics”.

The equity definitions offered by Gutiérrez, Cobb, and Hodge look at the effects or outcomes of equitable mathematics instruction. An opposing viewpoint would be to define equity as an “equitable” or equal distribution of mathematics instruction. This “one size fits all” approach was highlighted by Rousseau and Tate’s interview of a teacher. When asked how he views equity issues, the teacher responded by stating:

I try to make sure that I am working with all of my students equally or I am responding to students equally, and things like that...I think, I mean, being as open as possible with our students, answering as many questions for your students as you can, and treating them equally, is probably the biggest diversity issue as far as I am concerned. (Rousseau & Tate, 2003, p. 213)

The aforementioned teacher’s concern is with the process of the instruction, not the outcome or effects that it may have on certain groups of students. This type of approach is one where the race, culture, color, diversity of students is discounted or an attempt is made to look past it. Thompson (1998) offers a cautionary note of “colorblindness” within education where well-meaning White teachers often assume that it would stigmatize non-White students to notice their color. Thompson goes on to state that some teachers “bestow an honorary Whiteness” on all students. This bestowal further leading to and “assimilating the experience of people of color to that of Whites” within a culturally diverse society. Colorblindness fails to take into account the complex structures of race, culture, color, and class that are an integral part of our students’ lives and being. Not only does colorblindness limit teachers’ awareness of student characteristics and views, but fails to acknowledge the influences on schools, the impact of racism or its structures, and most importantly, makes it impossible to question or disrupt the ingrained

organizational structures of racism (Rousseau and Tate, 2003). Educators that adopt a colorblind approach deny students an “essential part of their being” (Lewis, 2003). Furthermore, Howard (2010) cautions that colorblind perspectives may contribute to internalized racism, reinforce racial hierarchies and student deficit models, and lead to the reproduction of racial and cultural hegemony in school structural decisions like curriculum choices, teacher expectations, testing procedures, and instructional practices.

A further operational definition for equity would be Ernest’s (2002) description of the three domains of empowering mathematics - mathematics, social, and epistemological. Mathematics empowerment allows students to gain “power” in the use of the language, skills, and practices of mathematics. Social empowerment helps students use mathematics as a tool for “sociopolitical” critique. Finally, epistemological empowerment provides students with the confidence to create and validate knowledge. By empowering students in these three domains, they are more equipped to level the racial, sex, and class imbalances that currently exists in advanced mathematics courses (Stinson, 2004).

One of the overarching goals of equitable mathematics instruction is that the knowledge it strives to transfer to students has “social, academic, and economic” currency (Bose & Remillard, 2010, p. 179). This is a similar view to Cobb and Hodge’s (2010) definition. This view runs counter to that of many teachers who feel that math is a “neutral” subject with no underlying cultural value. How can procedural algorithms provide anything other than ways to solve problems to students? As the past thirty years of educational research shows, the term “equity” is not only defined in different ways, but is also value laden. Similar, mathematics that is taught in school is also value laden.

“Math is Neutral” Viewpoint

A common view among educators is that mathematics is a culturally neutral subject (D'Ambrosio et al., 2013). Those who share in this view fail to understand the underlying social and cultural currents within mathematics education. Mathematics in today's society has an instrumental value as a means to obtain entry into college and high-earning careers and thus has extrinsic or structural significance (D'Amato, 1993). Gutiérrez (2007) describes mathematics instruction as being either dominant or critical. Dominant mathematics instruction reinforces the status quo of society, whereas critical mathematics instruction takes into account that students are situated within a society's class strata. Mathematics curriculum should reflect the lived realities of students otherwise we are likely to stereotype mathematics as knowledge that belongs to a few privileged groups (Secada, 1989).

While the results of both national and international tests show a widening between the mathematics achievement of certain groups, what is not so clear is why these gaps continue to exist. These “achievement gaps” can be looked at with a critical perspective. The entry point into critical theory and mathematics reform in our nation's public school is an often-quoted statement by Apple (2004), “...education was not a neutral enterprise, that by the very nature of the institution, the educator was involved, whether he or she was conscious of it or not, in a political act” (p. 1).

Pais and Valero (2010) state that the teaching of mathematics within schools perform two main functions. The first is to nurture the next generation of mathematicians and those who need mathematics to accomplish their work, thus assuring a strong global economy. This is reminiscent of Apple and Franklin's (2004) questioning “whether if

existing social and economic arrangements require that some people are relatively poor and unskilled and others are not?” (p.60). The second function is to ensure people enter society in which mathematics is seen as an indispensable requirement in being a citizen (p. 41). In this sense, mathematics performs a “gatekeeping role” in society.

The literature is clear that mathematics is not universally viewed as a “neutral” subject in school. If those within the education profession do not accept math as having both a cultural and social aspect to it, then how will they accept how mathematics is situated as a “gatekeeper” within our society? If elementary school students’ sex and/or ethnicity is predicative of their mathematics achievement, then a case can be made that the way mathematics is structured within elementary schools is not neutral, and “tips the scales” towards certain subgroups of students. Does ability grouping affect differently mathematics achievement of 4th grade students based on the sex and ethnicity of students?

Role of Math as a Gatekeeper in Society

Martin Luther King Jr. stated in November 1966 that the obstacles to the civil rights movement were “economic rather than legal, and tied much more closely to questions of class than to issues of race” (Jackson, 2013). Full citizenship and more importantly, economic access, is contingent on both mathematics and science literacy (Moses & Cobb, 2001). The importance of mathematics knowledge is manifest in both everyday activities and the employment arena (Stevens, Olivarez, Lan, & Tallent-Runnels, 2004). Schoenfeld (2002) argues that “purely” physical tasks of factory jobs have given way to technological tasks that require increased mathematics literacy.

A key to future success both academically and economically is successful completion of high school algebra because it provides a foundation on which higher-order mathematics, science, technology, and engineering courses are built (Evan, Gray, & Olchefske, 2006, p. 2 & 9). The U.S. Department of Education's National Math Advisory Panel's (2008) "Foundations for Success: The final report of the National Mathematics Advisory Panel" states that:

Success in mathematics instruction also is important for individual citizens, because it gives them college and career options, and it increases prospects for future income. A strong grounding in high school mathematics through Algebra II or higher correlates powerfully with access to college, graduation from college and earning in the top quartile of income from employment. The value of such preparation promises to be even greater in the future. The National Science Board indicates that the growth of jobs in the mathematics-intensive science and engineering workforce is outpacing overall job growth by 3:1. (National Mathematics Advisory Panel, 2008, p. xii)

In fact, the report goes on to state that completion of Algebra II has a significant correlation to success in college and future employment earnings. The college graduation rate for students who complete Algebra II in high school is twice that of those who do not (National Mathematics Advisory Panel, 2008, p. xiii). The outlook for students of differing ethnic and socioeconomic groups is bright if those students increase their success in high school mathematics courses (Evan, Gray, & Olchefske, 2006, p. 11). Carnevale and Desrochers (2003) assert that high school students who have completed Algebra II will be in the top half of earners (p. 26).

A longitudinal study that tracked high school and college students between 1980 and 1993 found that the level of mathematics studied in secondary school had the strongest influence on bachelor degree completion, and that these indicators are far more

intense for African-American and Latino students (Adelman, 1999). A 2006 revision of this study by Adelman (2006) found that:

By moving into the top two quintiles of the curriculum measure and completing a high school mathematics course beyond Algebra 2, African-American students who started out in a four-year college would hypothetically increase their bachelor's degree attainment rate from 45 percent to 73 percent; Latino students who did the same would hypothetically increase their bachelor's degree attainment rate from 61 percent to 79 percent. These increases were significantly greater than those for White and Asian students under this scenario, and, more importantly, were considerably greater than the influence of moving into the top two quintiles of either test scores or class rank/GPA. In other words, curriculum counts, particularly for minority students. (p. 5)

As has been shown by Adelman (1999, 2006), course taking has dramatic consequences for students as they move past the high school years and into college. The positive consequences of being in a more advanced course is more important than either test scores or GPA. These consequences are even more dramatic for students of color.

Mathematics holds a place within our society as a gatekeeper for both future academic and economic success. This gatekeeper status is even more pronounced for non-dominant ethnic, racial, and economic subgroups of students. The projected demographic shifts within our society will make for more a need to understand the role mathematics plays within school and classroom practices, and ultimately our students' lives.

Changing Demographics in U.S.

Economic disenfranchisement will only increase as our nation's demographics change over time. The U.S. Census Bureau estimates that by the year 2025, people of color will make up 38% of the population; by 2050 this estimate rises to 50% (El Nasser, 2004). With this change, U.S. schools and educators will need to be able to react to the changing demographics of our school age children. The changes in public school student

enrollments from 1993-94 to 2005-06 help to explain the increase awareness within academia of mathematics equity. During this period, White student enrollment in public schools remained at 28 million even though total student enrollments increased from 43 million to 48 million (Fry, 2007, p. 4). The percentage of overall White student enrollments decreased from 66.1% to 57.1% between 1993 and 2005. This same period showed increases of Hispanic enrollments from 12.7% to 19.8%, Black enrollments from 16.5% to 17.2%, and for Asian students 3.6% to 4.6%.

The future ethnic and racial makeup of the United States is expected to change at an increasing rate. Projections point to over half of the population belonging to a minority group (other than non-Hispanic White alone) by 2044 and by 2060 nearly one in five will be foreign born (Colby & Ortman, 2015, p. 1). The U.S. Census Bureau describes a major milestone is likely to be met in 2044. This is the year they project that the current majority group (non-Hispanic White) will make up only 44% of the population. This will be when our country becomes a “plurality” of racial and ethnic groups (Colby & Ortman, 2015, p. 9). Colby and Ortman (2015) go on to state the Hispanic population will increase from 17% of the population in 2014 to 29% of the population by 2060 (p. 9).

The way schools within this country adapt to changing demographics will be put to the test over the next quarter century. As the demographics shift, so will a need to focus more on the differences between the achievement gap and the opportunity gap.

Achievement vs. Opportunity Gap

National and international mathematics assessments give insight to the widening achievement gap in the U.S. Darling-Hammond (2010) states of all the nations that took the 2006 PISA (Program in International Student Assessment), the US ranked 21 out of

30 nations in science and 25 out of 30 nations in mathematics. Asian and White students from the U.S. scored above the PISA average while African American and Hispanic students from the U.S. scored below the average. Overall, Whites perform much better on these assessments than do Hispanics, who, in turn, do slightly better than African Americans. This is similar to the results of the 2012 PISA. The United States performed below average in mathematics in 2012 and is ranked 27th (PISA, 2012). Socio-economic disadvantage had a large impact on how the United States performed. 15% of the variation in student performance was explained by students' socio-economic status. This is in contrast with less than 10% in other similar countries. Even more alarming is the fact that in the U.S., as a student gets older, these gaps also keep increasing (Secada, 1992). The achievement gap between races and different ethnic groups has been the focus of educational researchers since the early-1990's. Secada's (1992) seminal work "Race, Ethnicity, Social Class, Languages, and Achievement in Mathematics" helps to shed light on the state of affairs of the "differentially effective" educational system within this country. Student's receive varying educational experiences based on social class, race, ethnicity, language background, sex, and other demographic characteristics (p. 623).

Do policy makers, educators, and researchers should continue to focus on the achievement gap in mathematics? For instance, data from the National Assessment of Educational Progress (NAEP) show that from 1990 to 2009, the achievement gap between fourth-grade Hispanic and White students remained at around 21-points. This gap increases to 26-points for eighth-grade Hispanic and Whites over the same time frame (Hemphill & Vanneman, 2011). NAEP data also show that an achievement gap is also persistent between Black and White students in schools where the density of Black

students is high. Additionally, the achievement gaps within each category of Black student density was smaller when student SES and other student, teacher, and school characteristics were accounted for (Bohrstedt, Kitmitto, Ogut, Sherman, & Chan, 2015). These data provide strong evidence that an achievement gap exists between certain races or more specifically that students who overall are Hispanic or Black perform much worse than students overall who are White. These disparities highlight the need for U.S. schools to address the mathematics needs of all students and to look at how poverty, race, and culture form a pedagogical nexus.

There is a growing amount of literature that examines equity based on educational opportunities and less on achievement. Gutiérrez and Dixon-Román believe that researchers place too much emphasis on the gap between the mathematics achievement of White, middle-class students and low-income and non-Asian minorities (2010). This “gap gazing” focuses more on maintaining the status quo from a dominant perspective and less on how students are situated within the social, political, and cultural framework. The primary criticism that Gutiérrez and Dixon-Román levy on researchers is that they stop at identifying the achievement gap and do not make substantial contributions to decreasing the gap.

Lubienski (2008) asserts that the term gap gazing is a determinant to mathematics instruction research. She states that “it does little to promote understanding of and mutual respect for diverse research methods and perspectives in mathematics instruction” (p. 350). Her definition of gap gazing is that it provides analyses of inequitable access to instructional resources and student outcomes and that it would be irresponsible for the mathematics instruction research community to pull back from such analyses. This

assertion is also held by Ladson-Billings (2006) who argues that "...an all-out focus on the achievement gap moves us toward short-term solutions that are unlikely to address the long-term underlying problem" (p.4).

Data show that opportunities to learn mathematics is still not equally distributed. Specifically, African American, Latino, and low-income students are frequently faced with inequitable distribution of mathematics opportunities (Flores, 2007, p. 29). Flores' focus is to minimize the amount of "gap gazing" to better understand the underlying causes of the achievement gap (p. 30). This switch moves researchers and educators alike to focus their attention on the opportunity gaps and less on the outcome focused achievement gap. Low-income, African American, and Latino students are less likely than White students to have qualified teachers, be exposed to rich contextual problems, and recommended for higher level mathematics. In fact, low track students are often assigned novice teachers or teachers who have fallen out of favor with school administrators (Secada, 1992, p. 646).

Not only are low-income and non-Asian minority students exposed to less qualified teachers and less rich curriculum, but they are underrepresented in advanced mathematics courses. Schools that enroll large numbers of low-income and non-Asian minority students tend to have more low-ability and fewer high-ability mathematics courses compared to schools that have more affluent demographics (Rousseau & Tate, 2003; Secada, 1992).

The shift from concentrating on the differences of achievement on high-stakes assessments to how educational opportunities are rationed will be challenging for many.

One of the ways that opportunities are rationed is by the equitable availability of high-quality teachers for all students.

Teacher Qualifications

Teacher qualification and experience is frequently cited as a symptom of the opportunity gap. Schools that have the highest poverty and those with the highest concentrations of minority students had nearly double the proportion of inexperienced teachers (Mayer, Mullens, & Moore, 2000; Wiener, 2006). The quality and experience of teachers matters not just in the U.S. but also across the 46 nations that participate in the “Trends in International Mathematics and Science Study” or “TIMMS.” The most significant predictors of mathematics achievement across all TIMMS participants is teacher’s certification, a major in mathematics, and at least 3 years of teaching experience (Darling-Hammond, 2010, p. 43). 67.6% of high-SES students were taught by these types of highly qualified/experienced teachers compared to only 53.2% of low-SES students, showing a 14.4% opportunity gap, which is significantly larger than the international average of 2.5% (Akiba, LeTendre, & Scribner, 2007, p. 378). Teacher practices are key in order for students to have an opportunity to learn mathematics. Teaching for mathematics proficiency and literacy is far more demanding than teaching procedural rules. High cognitive demand tasks require teachers to emphasize both meaning and thinking. These “opportunities-to-learn” dimensions are even more challenging for teachers when given the responsibility of students who have been traditionally marginalized by both educational and other social institutions (Bose & Remillard, 2010).

Some states have relied on alternative and “out-of-field” paths to teacher certifications in order to fill traditionally high vacancy teaching positions. Each of these

programs have varying qualifications, depending on the state and school district. Overall studies have found that the least prepared teacher recruits are disproportionately teaching predominantly low-income and minority students in central cities and poor rural areas (Flores, 2007, p. 32). Flores goes on to state these types of students are also more than likely to be taught by an out-of-field teacher, one who does not have at least a minor in the subject area.

The bottom line is that the proportion of teachers who are inexperienced, underprepared, or uncertified has a significant negative effect on student achievement after controlling for poverty and language background (Darling-Hammond, 2010, p. 50). A study by Jordan, Mendro, and Weerasinghe (1997) found students entering the third grade at the 50th percentile on a nationally normed mathematics test can end up at either the 27th percentile or 76th percentile over the next three years depending on the sequence of teachers (Wiener, 2006, p. 1326). A similar study conducted in 2006 found that teacher assignment was so significant that the difference between having a top-quartile teacher versus a bottom-quartile teacher for four years in a row was enough to entirely close the Black and-White achievement gap and had twice the impact as reducing class sizes from twenty-two to sixteen students (Gordon, Kane, & Staiger, 2006; Wiener, 2006). Between 12% and 14% of total variability in mathematics achievement gains within an elementary school year can be attributed to teacher differences (National Mathematics Advisory Panel, 2008, p. 35). A 2000 study by the National Council on Teacher Quality cautions that our country is on an educational treadmill in poor teacher education and low student results and that if “we” fail to teach mathematics well that we will only produce the “next crop” of weak students (Greenberg & Walsh, 2008).

There is a clear connection between teacher qualifications and the quality of mathematics instruction offered to student populations that have been underserved. Teacher qualifications are not the only component contributing to the mathematics opportunity gap. Another component is classroom practices and mathematics content that is presented to students.

Rich Contextual Problems

Not only are the more disadvantaged students being taught by the least qualified teachers, but they also receive lower quality instruction. Low-income and non-Asian minority students are frequently subjected to lower level mathematics instruction and basic skills concepts than higher-income and White students (Oakes, 1985). Ladson-Billings states that “sorting, grouping, or tracking students into lower levels where they receive minimal or no instruction from the teacher, while higher-track students are present with challenging and intellectually stimulating curriculum and instruction, is an example of how such inequity is structured.” (1995, p. 130). Mathematics has long been used to sort students, provide access to college, and ultimately filter people in to higher- and lower-wage work (Battey, 2013, p. 332). Roy (2000) goes on to assert that “equity will not be achieved if a disproportionate number of minorities remain unexposed to higher-level curriculum” (p. 41). The academic success of minority students can be better served with instructional approaches aimed at developing understanding, supporting peer interactions, and facilitating students’ participation than with those focused on rote memorization and decontextualized skill development (Remillard, 2000, p. 125).

The teaching of basic procedural skills has been referred to as a “pedagogy of poverty” (Habaerman, 1991; Bose & Remillard, 2010). The switch from basic

procedural skills to conceptual understanding and higher reasoning helps to prepare students for the demands of a technology-based, post-industrial economy. Development of a rich contextual mathematics curriculum, even in a low-income school, helps students to build on their existing knowledge base while providing for complex thinking and problem solving (Ladson-Billings, 1997).

Equitable mathematics instruction exists when all students are able to experience an epiphanic moment that can only come from rich contextual problems. Not only though do all students need these types of experiences, but they are also entitled to the opportunity to take advanced mathematics courses.

Access to Advanced Courses

Not only are low-income and non-Asian minority students exposed to less qualified teachers and less rich curriculum, but they are underrepresented in advanced mathematics courses. Schools that enroll large numbers of low-income and non-Asian minority students tend to have more low-ability and fewer high-ability mathematics courses compared to schools that have more affluent demographics (Rousseau & Tate, 2003; Secada, 1992; Silver, Smith, & Nelson, 1995). Mathematics has long been used to sort students, provide access to college, and ultimately filter people in to higher- and lower-wage work (Battey, 2013, p. 332).

A 2008 report published by the U.S. Department of Education titled “Foundations for Success” recommends that “all” students been provided opportunities to enroll in an algebra course by Grade 8 (National Mathematics Advisory Panel, 2008, p. xviii). This is significant given the linkage between completion of Algebra I within middle school and future academic and employment success.

The opportunity gap manifests itself through the lack of pathways available to all students to take advanced mathematics courses. When advanced mathematics is rationed, for whatever reason, the nation as a whole starts to take on an “educational debt” that is analogous to its fiscal debt.

Educational Debt

Ladson-Billings (2006) describes the achievement gap that has historically existed in this country metaphorically. She states that “our” (researchers, politicians, media, etc...) focus on the achievement gap is likened to the focus on the budget deficit, but that it should be more likened to the national debt. The national debt has been the result of year-after-year of deficit spending. This metaphor shifts the focus to the accumulating damage that an “education debt” plays on our society both collectively and on an individual basis. Instead of investing in the resources needed to school low income and minority students, the education debt leads to a slew of social problems such as crime, low productivity, low wages, and low labor force participation. She states that the education debt is the result of the funding disparities between schools serving White students and those serving students of color. Another factor is that communities of color have less representation in the educational decision-making processes. The importance of addressing the educational debt is threefold. The first is the impact the debt has on the current education process. The second is the ability to understand how the debt is related to past educational research findings. The third is the potential to create a better future for all of our students.

Any study that attempts to look at the equitable distribution of mathematics within a school setting must include an operational definition of equity. By viewing equity in a

way that can enhance the future academic and economic opportunities of students, teachers are acknowledging the “non-neutral” way mathematics inhabit towards society. This acknowledgement of the ability of mathematics to change the lives of students give them, to use Bruner’s (1986) term, clout. This is increasingly important as the demographics of the U.S. shift to a more culturally diverse society.

Thirty years of research into the equitable distribution of mathematics instruction has yielded varying operationalized definitions of the term “equity.” The underlying attribute of these definitions is that mathematics instruction has the ability to affect the lives of students, in either a positive or negative way. Social justice is an often-overlooked goal when mathematics curriculum and classroom practices are revised. As teachers, do fully understand that such emblematic school processes as sorting students into homogenous mathematics groups could lead to the unequitable distribution of mathematics instruction?

Social Justice

Paulo Freire originally wrote in 1968 that “the oppressed and those who are in solidarity with them will not gain liberation by chance, but through their quest of it and the necessity to fight for it” (Freire, 2000, p. 45). One would hope that educational researchers and teachers would consider themselves in solidarity with the oppressed. It is with this sobering thought that mathematics equity literature was explored through a social justice lens. Literature was explored into following functional areas: 1) critical perspective, 2) mathematics as a way to challenge authority, 3) valorization of mathematics, 4) educational funds of knowledge, 5) mathematics communities of practice, and finally 6) tracking, grouping, and sorting within mathematics classrooms.

Critical Perspective

While the results of both national and international tests show a widening between the mathematics achievement of certain groups, what is not so clear is why these gaps continue to exist. These “achievement gaps” can be viewed through a critical lens. The entry point into critical theory and mathematics reform in our nation’s public school is an often-quoted statement by Apple (2004), “...education was not a neutral enterprise, that by the very nature of the institution, the educator was involved, whether he or she was conscious of it or not, in a political act” (p. 1).

This implies a person must comprehend the notion that schools are under the hegemonic influences of a dominant ideology. And if this is true we need to reform the system in a positive and equitable manner to improve achievement for all races, ethnicities and social classes. This comprehension gives insight into Apple’s statement about education being a political, non-neutral pursuit. To extend this idea of the hegemony in our schools one can look to Althusser’s definition and explanation of the “Ideological State Apparatus”. He writes that the “Church has been replaced today in its role as the dominant Ideological State Apparatus by the School...the Apparatus playing a determinant part in the reproduction of the relations of production of a mode of production threatened in its existence by the world class struggle” (Althusser, 2006). At this moment in history, schools portray themselves as the instrument of social mobility, but in reality, they continue to perpetuate the existing social orders. Pais and Valero (2010) further explore the role of a dominate ideology in mathematics equity research. They assert that the framing of mathematics instruction within the realm of future employment and success “conceals” the notion of what it means to be a worker and

citizen in a capitalist society (p. 41). Žižek continues the criticism of mathematics research by focusing on the National Council of Teachers of Mathematics (2000) standards from 2000. He states that the standards express “an official discourse” of the virtues and democratic goals that society stand for, but will fail in its implementation (Žižek, 2009). Pais and Valero help to further this idea by explaining that a more democratic discourse is needed to ensure that mathematics instruction provides for other roles than what is currently expected (Pais & Valero, 2010).

Mathematics curriculum and instructional practices frequently comes under the microscope and is modified to help improve assessment scores. These changes habitually address the achievement gap, but overlook the opportunity gap. Critical pedagogues focus is not to modify, but to transform and revolutionize the structural framework of mathematics instruction. The lives of students can then be transformed. One way that equitable mathematics instruction can change the lives of students is to give them the tools, confidence, and clout to challenge authority.

Using Math to Challenge Authority

A shift from dominant to critical mathematics instruction is realized when the cultural identities of the students are used to help them to address social and political issues (Gutiérrez, 2007). Gutiérrez (2007) further claims that for mathematics to be taught through a critical lens that students should be able to recognize the relationship between mathematics and power. Additionally, they should be able to investigate and question the knowledge base of the mathematics curriculum (p. 46-47).

Freire (2000) cautions against education that is based upon teachers dispensing (banking) knowledge into passive recipients to be passively retrieved at a later time. He

argues that “the more completely they [students] accept the passive role imposed on them, the more they tend simply to adapt to the world as it is...” (p. 54). Gutstein (2007) takes Freire’s banking model and further expands it within a questioning model of mathematics pedagogy. His focus on social justice asserts that critical mathematics pedagogy should include a sociopolitical consciousness and a sense of social agency (p. 54). The top layer of Nunley’s three-layered curriculum asks students to ponder ethical, moral, and/or belief-driven questions (Nunley, 2004). It is through these types of curricular choice that teachers can situate mathematics that give students a social and cultural empowerment.

The complex interactions that led to the assimilation of mathematics knowledge is the focus of Martin’s (2007) research. He describes mathematics socialization as the experiences that individuals and groups have within different contexts that legitimize participation in mathematics. Mathematics identity is described as the dispositions and beliefs that individuals develop regarding their ability to participate in and accomplish mathematics tasks and to use mathematics knowledge to change the conditions of their lives (p. 150). How students construct identity is further expanded by Cornell and Hartmann (2006) that identify three constructs: boundary, perceived position and meaning. The boundary construct distinguishes one member of a group from another (e.g. gifted and not-gifted classrooms). Students are aware of these boundaries from an early age. They know who are in the low and high reading groups. Perceived position is what how a member or group’s status is located within the stratified levels or positions of power. Not only do students know where the boundaries are, they know how each group is positioned in relation to the others. Not only do students construct their membership

within groups and understand how those groups are positioned, but the construct what membership in these groups mean.

Mathematics instruction can be a transformative influence in the lives of students of all ages. It can empower students with the ability to challenge authority in a manner that can have profound impacts on their lives. One way in which this can be accomplished is through the acknowledgement there is value in the way mathematics is used in the lives of students at home. Educators need to recognize and accept that students walk through our classroom doors with existing mathematics beliefs and practices.

Valorization of Mathematics

Issues of equity and ultimately social justice cannot be addressed without discussion of how societies, communities of practice, and individuals assign value to mathematics knowledge. Abreu's (2007) concept of social valorization stemmed from research within communities of Brazilian farmers. Distinct forms of mathematics associated with different communities of practice within these rural societies were assigned varying status by adults and children. Abreu (1995) noted that there was a complex relationship between the farmers' traditional mathematics practices and those learned within school. Within the social context, farming mathematics was assigned a lower status, whereas school mathematics was higher. Abreu's research into social valorization brings to light the cultural aspect of mathematics instruction. Frequently it is the teacher who decides the value of the mathematics being taught, without regard to the student's background. Abreu (2007) stated that, "for the teacher, farming mathematics was associated with a type of person the children did not want to become." (p. 120)

A key aspect of Abreu's (2007) work with social valorization is that it positions individuals and groups to one another in the broader societal context. Additionally, it influences the ways individuals and groups define who they are. Further questions explored by Abreu include how the differential valorization of students' home experiences impact their development of their academic experiences, do children exposed to varying mathematics practices gain understanding about the underlying valorization of these practices, and how do teachers' views and valorization of students' home background influence their teaching practices? The exclusion of the mathematics knowledge of minority groups has the consequence of building negative self-esteem and cultural identity, which may ultimately lead to the inability to access the construction of mathematics meanings (Abreu & Cline, 2003, p. 13). This is reiterated by Valero (2009) who stated that mathematics instruction can be defined as a series of social practices and that these practices are to be found not only within the classroom but within family and local practices (p. 69). Students from minority groups are either empowered or disabled as a result of how they interact with their teachers. The result of these interactions is that there is a need to understand the perceptions of learning opportunities from both the perspective of students and teachers if the mathematics underachievement of certain groups is to be understood (Planas & Civil, 2009, p. 393).

Mathematics has value that is often determined by those whose lives it situated in. In other words, the mathematics that is of value to our students is the mathematics that they use within their daily lives. This is a shift in thinking from the traditional viewpoint that the only mathematics that is important is what is presented within the classroom. One

way to make this paradigm shift is for educators to acknowledge the extensive and valuable funds of knowledge that students bring to the classroom from their homes.

Funds of Knowledge

Frequently teachers assign value to mathematics based on the view that these same teachers have of their students' home lives. The accumulated knowledge that exist within homes is often referred to as funds of knowledge (Velez-Ibanez & Greenberg, 1992). These funds of knowledge help educators to make connections between the existing background knowledge of students and mathematics concepts. Another key aspect is that by acknowledging funds of knowledge that exist within a student's home, a teacher can help that student find value with mathematics that is situated within their life. This shift helps to move how mathematics is situated from the life of the teacher to that of the student. Rich (1979) wrote that, "when someone with the authority of a teacher, say, describes the world and you are not in it, there is a moment of psychic disequilibrium, as if you looked into a mirror and saw nothing." Knowledge is gained as people interact with their worlds through artifacts, practices, and the use of language (both written and oral) (González, Andrade, Civil, & Moll, 2001, p. 121). The concepts studied in school form the conceptual fabric for the development of new knowledge as they are incorporated into the existing scheme.

By recognizing that there are funds of knowledge that exist within even what the casual observer would consider a pathological home life, a teacher can facilitate a student assigning value to mathematics. In this way, zones of practice can provide an invitation to students into a world in which in-school and outside-school mathematics can become one (González, Andrade, Civil, & Moll, 2001, p. 128). By blurring the lines between in-

school and outside-school mathematics, an educator can build a community of practice within their classroom.

Communities of Practice

Research into the role that cultural plays in mathematics classroom have started to focus on how communities of practice are formed within them. A functioning classroom community between a teacher and students is shaped by norms and practices that include issues of competence, ownership, and alignment in engaging in the community (Bartell, 2007). Cobb and Hodge (2002) state that the way that students participate in these classroom communities of practice are affected by the cultural diversity within the community. A relational perspective conceptualizes students' participation in the practices of local, home or broader communities and orient research to their lives outside the classroom. Without this relational perspective, some students' mathematics reasoning associated with particular groups are frequently treated as illegitimate in school (p. 254). Wenger (1998) states these communities of practice are not static but evolve into forms of mutual engagement. Additionally, communities of practice fine tune their enterprise to hold all accountable. Finally, they develop organic styles and discourses. Within this structure, community members' learning occurs when they participate in and contribute to the evolution of the group's norms and practices. Through this community of practice lens, students are seen as legitimate peripheral participants within their home communities in which they are valued as participants (Cobb & Hodge, 2002).

As students become legitimate community participants within a mathematics classroom and gain new skills and knowledge they develop new identities relative to their community (Nasir, 2002). This identification process can act as a motivator for new

learning (p. 240). The cautionary note for teachers is that students of dominant groups could already be seen as having an identity acceptable to the norms and practices of the mathematics classroom community and are rewarded for this identity (Lubienski, 2002).

Cobb and Hodge (2002) use three different aspects of the classroom micro-culture within their relational perspective: 1) social norms, 2) socio-mathematics norms, and 3) classroom mathematics practices. Social norms include ways that students were expected to explain and justify their methods of solving a problem, making sense of explanations given by others, and questioning alternatives. Socio-mathematics norms are focused on classroom interactions that are specific to mathematics. They include efficiency within solutions, legitimizing solutions, and acceptability of mathematics explanations.

Classroom mathematics practices focus on ways of reasoning, arguing, and symbolizing mathematics ideas (p. 269). Cobb and Hodge (2002) argue that by looking at mathematics classrooms through a relational community of practice lens, students' local and home communities can be taken into account. They refer to the collective classroom micro-culture as the cultural capital of mathematics. Their main thesis being that "how students participate in the continual regeneration of this cultural capital is of fundamental importance as it relates directly both to their access to opportunities to develop forms of mathematics reasoning that have clout and to the identities that they construct as doers of mathematics." (p. 271)

Building a community of practice within a mathematics classroom is a complex process that must account for social and academic forces. Similarly, the process and effects of tracking, grouping, and sorting of mathematics is complex.

Tracking, Grouping, and Sorting

One possible structure attributed to the widening achievement gap is that of homogenous grouping and tracking of students. Increased time on mathematics and the taking of advanced coursework are two of the most powerful predictors of student achievement (Secada, 1992). Over twenty-five years ago Oakes (1990) stated:

Some see schools as meritocratic institutions that consider achievement itself as the principal mediator of opportunity, arguing that children who achieve more are better able to benefit from and more deserving of the limited resources that are available. Others explain opportunity, achievement, and participation as a function of mental capacity; for them, the most important opportunities are conferred at birth or before. (p. iii)

Oakes statement is a sobering thought. If it is to be believed, then achievement begets achievement. Those who exhibited high achievement are given more educational opportunities.

Grouping and tracking students inevitably separates students by race, ethnicity, native language, and class (Loveless & Diperna, 2000). In the past, Latino students who scored within the 60th percentile were 50 percent less likely to be placed into a college preparatory class when compared to White and Asian students (Oakes, 1992). Not only are non-Asian minority and poor students tracked into lower mathematics courses, but those courses tend to have less qualified teachers (Secada, 1992). This trend is still true today, with Black, Hispanic and poor children still frequently tracked into remedial classes, while middle-class White children are tracked into honors courses (Loveless & Diperna, 2000).

Tracking is put into place because educators believe and argue that it better meets the individual needs of students. The antecedent of tracking students into low-ability mathematics classes at the secondary level typically begins with recommendations

initiated by the student's upper elementary school teacher. Additionally, there is the belief that students are better served later in life by putting students into homogenous mathematics groups. Additionally, there is widespread belief that children's intellectual difference is so great and that there is such a perceived varying ability levels that they need to be taught in separate classrooms (Oakes, 1987 & 1990).

Typically, students are grouped based on judgements of their ability and educability (Secada, 1992). Secada goes on to state that there is evidence that schools do not sort students only on merit and they do not sort students uniformly. There are substantial ethnic differences in how students are assigned to tracks, even after adjusting for prior achievement.

Hattie's meta-analyses (2008) has found that tracking not only has minimal effects on learning outcomes but has profound negative equity effects. The resistance to elimination of tracking programs is typically from "high-track" teachers and parents who have benefited from tracking (Mathis, 2013). The goal of curricular stratification (tracking) in American schools has been to create homogeneity based on student ability at the classroom level (Burriss, Welner, and Bezoza, 2009). Research has shown that tracked classes have a wide range of student ability levels due to how enrollment decisions are made. These decisions include not only test scores and prior achievement, but student behavior, parent preference, completion of prerequisites, teacher judgment, and counselor guidance (Useem, 1992). Not only is the selection process for tracking students suspect, but teacher qualification in low-track classes is a concern. Low-track courses frequently had more out-of-field teachers than do high-track courses (Tate, 2008). Stinson (2004) gives a qualitative insight as to what tracking looks like at the classroom level.

My experiences as a secondary mathematics teacher, preservice-teacher supervisor, and researcher supported Oakes's (1985) assertions that often students are distributed into "ability" groups based on their race, gender, and class. Nonetheless, my perception after five observations was that ability grouping according to these attributes was diminishing—at least in these elementary schools. In other words, the student make-up of each mathematics lesson that I observed appeared to be representative of the demographics of the school.

However, on my sixth observation, at an elementary school with 35.8 % Black, 12.8 % Asian, 5.3 % Hispanic, 3.5 % Multi-racial, and 0.5 % American Indian children, I observed a 3rd grade mathematics lesson that was 94.4% White (at least it was 50% female). The make-up of the classroom was not initially unrepresentative of the school's racial/ethnic demographics, but became so shortly before the start of the mathematics lesson as some students left the classroom while others entered. When I questioned why the students were exchanged between classrooms, I was informed that the mathematics lesson was for the "advanced" third graders. (p.8)

Is this typical of most schools? How were students selected for the advanced mathematics class?

It is hard to uncouple both the causes and effects of the opportunity gap in mathematics without addressing the social justice aspects of them. Critical pedagogies like Apple (2004), Althusser (2006), Bourdieu (1977), Freire (2000), and Žižek (2009) have cautioned that without change, schools will continue to replicate the same inequalities of they have existed since the industrial revolution. It is under this social justice banner that students can use mathematics to help challenge the status quo of economic and social equalities. These funds of knowledge help to shape communities of common practice within the classroom. Only through this lens can the instructional structures of tracking, grouping, and sorting be addressed within schools.

Teacher Judgement

The final theoretical lens that mathematics equity was explored within this literature view was that of teacher judgement. Educators at all levels are asked on a daily basis to make judgements on their students' performance, behavior, and work habits. These judgements all also used to help plan instruction and guide school practices. Frequently teachers are asked to make decisions on which students will have access to advanced mathematics opportunities. This section will explore the literature related to teachers' judgement in the form of evaluations and naturalistic expectations.

Teachers Evaluations

Teacher's perception of a class's ability plays an important role in deciding what to teach and how to teach it. Judgments about academic ability often lead to the segregation of students into separate mathematics classes and to enrollment in different senior high school courses (Oakes, 1990, p. 17). Compounding the detrimental effects of tracking is that teachers are unable to accurately determine which students are qualified for or would benefit from more advanced mathematics courses. Factors such as SES, parent influence, and racial/cultural membership are frequently used as sorting mechanisms (Darling-Hammond, 2010; Oakes, 1990; Secada, 1992). Teacher judgments are frequently used for instructional planning, screening, placement, and referrals (Martinez, Stecher, & Borko, 2009).

There is currently a lack of research into the structural processes that upper elementary school teachers use to recommend students for advanced mathematics courses and what the effects of this grouping has on students' mathematical achievement. Students moving into the middle school grades (6 through 8) are often placed within

homogeneous mathematics groups based on several factors, one of which is recommendations made by their previous elementary school mathematics teacher. The factors that elementary teachers use to make these recommendations may vary from perceptions of students' work ethic, demographics (i.e. sex, race, cultural identity, SES, etc...), mathematics ability, and student home life.

Martinez, Stecher, and Borko (2009) assert there are two contrary held beliefs regarding teacher judgements of student achievement. The first is that teachers are able to assess student achievement with a high degree of accuracy and validity. The second belief is that teachers are not always able to distinguish accurately between achievement and traits such as motivation, engagement, and measured or perceived ability. Their research found the strength of the correlation between teacher achievement ratings and standardized test scores differed substantially from classroom to classroom (Martinez, Stecher, and Borko, 2009, p. 92). Teachers are not uniform in their ability to accurately predict student success.

Seceda (1992) stated that teachers uniformly apply judgements across varying student groups in a fair manner. He goes on to describe how these judgements are "constructed" in a biased manner. Teachers are accurate in achievement (i.e. standardized tests) but not successful in predicting learning (i.e. algebraic reasoning). Seceda contended that the generalized notion that teachers' have of their students' educability seemed based on a combination of factors like classroom behaviors, along with ability to learn.

The ability of teachers to accurately judge and predicate student achievement and future success is an area intricately tied to mathematics equity. Naturalistically formed teacher expectations is one area of research into teacher judgements.

Naturalistic Expectations

Teachers can form judgements and expectations of student through interactions, reviews of school records, and past knowledge of siblings and family members (Strutchens, 2000, p. 8). Strutchens go on to state that studies of naturalistically formed expectations have found that teachers treat students differently based on perceptions of students' potential. These biases can be related to ethnic, racial, and/or social-economic groupings. Citing a 1993 study by Irvine and York (1993), Strutchens found teachers stated the primary reason that African American students fail was due to lack of parental support. Additionally, she found teachers stated the primary reason that Vietnamese and Hispanic students fail is due to language difficulty.

The power of teacher expectations has been studied for close to a half century. Rosenthal & Jacobson's (1968) groundbreaking study, "Pygmalion in the classroom: Teacher expectations and student intellectual development" may have raised more questions than answers regarding the affect that teacher expectations have on student achievement. Jussim and Harber's 2005 review of Rosenthal and Jacobson's original study tried to shed light on the actual relationship between teacher expectations and achievement. At face value, Rosenthal & Jacobson's 1968 study found that teacher expectations created a self-fulfilling prophecy in first through sixth grade students (Jussim & Harber, 2005, p. 133). At a closer look, the effect of teacher expectations on student achievement was statically minimal. One reason for this is the accuracy of teacher

judgements; with accuracy referring to teacher expectations predicting but not causing student achievement. In fact, as accuracy increases, the potential for self-fulfilling prophecies declines and vice versa (p. 138).

Twenty-five years of research into teacher judgements has left confounding views on the efficacy of teacher judgements. On one hand teachers tend to be a good judge of future success on standardized achievement test. On the other hand, teachers are frequently unable to determine which students would benefit the most from advanced mathematics courses (Walter G Secada, 1992). This is a critical disconnect giving the fact that teacher judgements drive classroom practices like curriculum choices, pacing, and recommendations for future mathematics courses.

Teacher judgements about student achievement and future success are a key component of mathematics equity. For almost a half-century, educational researchers have been trying to piece together all of the factors that teachers use to make classroom decisions.

Guiding Theories

The focus of this literature review has been in the area of the mathematics equity. Past and current research was viewed through three lenses: 1) mathematics equity, 2) social justice, and 3) teacher judgements. Each of these lenses were used to guide this study's research question into how ability grouping differently effects the mathematics achievement of 4th grade students based on their sex and ethnicity?

The theoretical framework of primary interest based on the issue of equitable distribution of mathematics knowledge is that of situated learning within communities of practice. Situated learning is concerned with the relational "interdependency" between

agent (learner) and the world (Lave, 1991). Meaning is socially negotiated within a socially and culturally structured world. Furthermore, the construction of meaning is contingent on the action of persons engaged within the activity. Lave (1991) further asserts that the development of an identity within a community are part of the same process as becoming skillfully knowledgeable (p. 65).

The relationship between membership within a community of practice and acquisition of new knowledge is complex. Nasir (2002) states that “as members of communities of practice experience changing identities, they come to learn new skills, facilitating new ways of participating, which in turn, helps to create new identities relative to their community” (p. 239). Additionally, “identities can act as a motivator for new learning, prompting practice participants to seek out and gain the new skills they need to participate in their practice more effectively” (p. 240).

The equitable distribution of mathematics knowledge within communities of practice (i.e. mathematics classrooms) can be seen by critical theorists as helping to reproduce the ideological hegemony of the dominant class. This can occur if only students from the dominant group are seen as contributors to the discourse of the classroom. Bourdieu (1977) argued that the ways that students from dominant groups talk are generally deemed more acceptable than other students and are rewarded as such.

Cobb and Hodge (2002) assert that socio-mathematics norms are as important within mathematics classrooms as are social norms. They assert that:

These norms can include what counts as a different mathematics solution, a sophisticated mathematics solution, an efficient mathematics solution, and an acceptable mathematics explanation. The last of these norms has proven to be particularly important in a number of empirical analysis that we have conducted, and we anticipate that it might also be highly relevant to investigations of cultural diversity and equity in mathematics instruction. (p. 268)

They go on to further state that:

This view of cultural diversity and equity in mathematics instruction stems from the fact that it focuses on communal process and simultaneously takes mathematics ideas seriously. Crucially, it challenges the traditional metaphor of mathematics as objectified content that is placed in the container of the curriculum and is instead consistent with the view of mathematics as a socially and culturally situated human activity. (p. 270)

The assertion that mathematics instruction can be subjective versus objective is a radical departure for traditional thought.

Issues such as competence, ownership, and alignment with community engagement are at the forefront of promotion of a mathematics classroom into a functioning community (Lester, 2007). These areas are shaped by both teachers and students. Lester continues to state that by recognizing and valuing the multiple ways that students participate in mathematics can empower and include students that are otherwise “sidelined” (p. 408).

Ernest (2002) makes the argument that there are several domains of empowerment within a mathematics classroom. They are mathematics, social, epistemological, and professional empowerment of mathematics teacher. Mathematics empowerment concerns the gaining of power over language, skills and practices of using and applying mathematics. Social empowerment concerns the ability use mathematics to better one’s life. Ernest maintains that true social empowerment goes beyond opening doors of opportunity though successful completion of high stake test (p. 4). True social empowerment within mathematics includes a critical understanding of how to identify, interpret, evaluate and critique the mathematics embedded in social, commercial and political systems. Epistemological empowerment is the personal sense of power over the creation and validation of knowledge. Ernest uses Belenky’s (1986) epistemological

powers of the individual model (see Table 1: Belenky's (1986) Epistemological Powers of the Individual) as a lens to look at epistemological empowerment within a mathematics classroom. The final stage of empowerment being the construction of knowledge within the learner. Finally, the professional empowerment of mathematics teachers concerns the transformation into an autonomous and reflective teacher with the confidence to construct and critically assess teaching, learning, and assessment.

Table 1: Belenky's (1986) Epistemological Powers of the Individual

Silence	
accepting the pronouncements of authority passively	
Received knowledge: The Voice of Others	
accepting the pronouncements of authority, but able to repeat them	
Subjective knowledge: The Inner Voice	
responding intuitively, and valuing their own subjective intuitive judgements	
Procedural Knowledge	
seeking objective knowledge by means of either Separated or Connected Knowing	
Separated Knowing (impersonal, critical and rational reasoning, aiming at justification and proof), or	Connected Knowing (arising from experience of relationships, and requiring intimacy and equality between self and object, knower and known, aiming at uncritical, empathic understanding)
Constructed Knowledge: Integrating the Voices	
all knowledge is understood to be constructed by the knower herself; and the voices of intuition and of reason are integrated into a way of knowing that depends on the frame of reference of the knower as well as on the overall context; the knower is put into relation with the known.	

The guiding theories of situated learning, communities of practice, and empowerment are interwoven into a common framework. This framework allows mathematics teachers at all levels the knowledge and permission to build a mathematics

community of practice within their classroom. This community of practice can empower students to use and build upon the mathematics values that they bring from home.

Students can then take this empowerment with them outside the walls of the classroom as they interact with the real world.

Lack of Current Research

The field of educational research focused on mathematics equity has seen increased scrutiny over the past twenty years. Secada, Fennema, and Adajian (1995) warned that the urgency of researchers to look into issues of equity has led to a rush for answers and solution (p. 149). Lubienski (2002) argued that “although researchers and reformers give attention to equity, such work tends to ignore relevant social and cultural issues” (p, 103). A past review of mathematics research shows limited emphasis on classroom processes and ethnicity, class, and sex in relation to student cognition (Lubienski, 2002, p. 107). Lubienski states that most mathematics instruction researchers are unfamiliar with theories relating to culture and power. A background in anthropology and sociology is just as important as that of mathematics pedagogy. Mathematics instruction research with a diversity focus is both marginalized and underdeveloped (Cobb & Hodge, 2002, p. 250). Additional research is needed to explore the systematic and structural aspects of inequity in mathematics instruction (Lester, 2007, p. 411).

What Needs to be Studied

The field of educational research viewed through a cultural diversity and equity lens is filled with opportunities. Almost twenty-five years ago Seceda (1992) stated that there is evidence that schools do not sort students strictly on the basis of merit, nor are they sorted uniformly (p. 646). He based this statement on the sorting and tracking of

secondary students. Is this the case with how mathematics knowledge is rationed to 4th, 5th, and/or 6th grade students? Has research neglected to determine how the mathematics opportunity gaps are initiated and galvanized in the upper elementary school grades? Building on Secada's work, Cobb & Hodge's (2002) attempted to look at the linkage between the building of community of practices and mathematics thinking within the classroom. What is not clear is whether or not there is a correlation between elementary school mathematics teachers' ability to build communities of practices and how this shapes their views on the equitable distribution of advance mathematics opportunities. This question goes hand-in-hand with Abreu and Cline's (2003) work on social valorization of mathematics practices. They questioned how teachers' views and valorizations of students' home backgrounds and practices influenced both school practices and issues of equity (p. 124).

Research Question

A common thread found throughout this review of literature related to mathematics equity is that mathematics is not equally rationed across all subgroups of student populations and that the effects of sorting students into homogenous ability groups is not fully understood. The literature is clear that students are sorted into low, grade-level, and advanced mathematics courses throughout their secondary schooling. It is not clear within the literature whether or not students are sorted in elementary school and what the effects are of this sorting if it exists.

The purpose of this literature review was to explore mathematics equity as it applies to elementary mathematics students. A lack of research was found in the area of not only the process of sorting elementary students but also the effects of sorting on

mathematics achievement. Of primary interest from the preceding review is how does this sorting take place within elementary schools. The way students are sorted into varying mathematics groups gets to the heart of the equitable distribution of mathematics instruction. Furthermore, the effects of sorting help to make a connection between the current lives of our students and what opportunities they will have in the future.

The equitable distribution of mathematics instruction and the effects of sorting are key to this study's two research questions.

Research Question 1 - Does ability grouping affect mathematics achievement of 4th grade students?

Research Question 2 - Does ability grouping affect the mathematics achievement of 4th grade students differently based on sex and ethnicity of students?

CHAPTER 3: METHODS AND PROCEDURES

Purpose

The purpose of this quantitative study is to compare the mathematics achievement of 4th grade elementary school students who are grouped in either homogenous or heterogeneous mathematics classes. Student achievement was measured from the beginning of 4th grade till the 4th quarter of the school year based on the STAR (Learning, 2014) standardized mathematics assessment. Independent samples t-test were used to analyze variations in growth based on whether or not students were grouped and whether students were male or female. Analysis of variance (ANOVA) and a linear regression were used to analyze predictability in growth based on the level of mathematics ability grouping and by student ethnicity.

There were two guiding research questions and associated hypothesis for this study.

Research Question 1 - Does ability grouping affect mathematics achievement of 4th grade students?

H₀: Ability grouping does not have an affect on mathematics achievement of 4th grade students.

H₁: Ability grouping does have an affect on mathematics achievement of 4th grade students.

Research Question 2 - Does ability grouping affect the mathematics achievement of 4th grade students differently based on sex and ethnicity of students?

H₀: Ability grouping by mathematics achievement of 4th grade students does not affect students differently by sex and ethnicity.

H₁: Ability grouping by mathematics achievement of 4th grade students affects students differently by sex and ethnicity.

This chapter describes the research design, setting, sampling, instrumentation, timeline, data collection and analysis.

Research Design

This study used a causal-comparative design in an attempt to find relationships between independent and dependent variables (see Appendix B). A causal-comparative design “seeks to find relationships between independent and dependent variables after an action or event has already occurred” (Salkind, 2010). The independent variables were student sex, student ethnicity, whether or not students were grouped into mathematics ability groups, and the level of mathematics ability group. The dependent variable of this study was the overall change in mathematics achievement from fall to spring. The goal of this quantitative study is to determine whether the independent variable is related to the outcome variable by comparing two or more groups of individuals after the event has occurred (*ex post facto*).

A causal-comparative design is frequently used within educational research where independent variables are already fixed in place (Salkind, 2010). Classrooms are not re-organized based on the experimental design constraints of a research study. Another key aspect to causal-comparative designs is subjects are already in groups in contrast to a true experimental research design where subjects are randomly selected (Salkind, 2010).

Setting and Sample

Population Definition

Within Idaho, there are 23,000 4th graders enrolled in 368 schools. The population of this study consisted of 4th grade students from within a school district in Idaho. As of the 2014-2015 school year, this district had approximately 2,000 4th grade students in 32 elementary schools. One school district was selected for this study. There were three primary reasons this study was limited to the one school district. The first was that this district had existing research committee procedures that allowed for the timely approval to begin this study. The second reason was that this district had a diverse student population of 4th graders within its 32 elementary schools. Finally, this district had already administered both fall and spring STAR mathematics assessments to its 4th graders. This would limit the disruptions to students because classroom teachers would not have to administer another assessment.

Sampling Method

Nonprobability convenience sampling was used to acquire volunteer teacher participation from the one school district. McMillan (1996) states that though convenience sampling should be used cautiously due to its lack of generalizability, it may be the only sampling method possible in an educational setting, especially when the primary purpose of the study is to better understand the relationships that exist. Dörnyei (2007) asserts that members of the target population are selected for the purpose of a study if they meet certain criteria like ease of accessibility or a willingness to volunteer. The target population was 4th grade teachers within the school district included in this study. Teachers agreed to take an online questionnaire in order to participate. By agreeing

to participate, teachers also gave consent to the collection and analysis of their students' mathematics achievement scores.

Participants

Teachers were recruited from a mid-size school district in southwest Idaho. According to the district's webpage, it has approximately 26,000 kindergarten through twelfth grade students in 52 schools. Roughly 14% of the district's students come from a non-English speaking background, with 47% eligible for free or reduced-price lunch. The average class size for 4th grade is 32 students per class. The district's overall student population is 78% White, 11% Hispanic/Latino, 4% Asian and 3% Black.

School district approval was granted via the district's research committee procedures. This approval was granted by email within twenty-four hours of submission. The school district's approval allowed direct recruitment of 4th grade teachers without having to ask permission of the respective principals. The school district agreed to provide an email roster of all 4th grade teachers within the district and consolidated student growth data. All district principals of elementary schools were emailed by the district mathematics supervisor to make them aware that this study was given district approval.

The participants for this study were drawn from 4th grade teachers and their respective mathematics students. Emails were sent to all eighty-five 4th grade teachers within the district to secure their voluntary participation with the study. Twenty-seven teachers consented to participate with the study (31.7% participation rate). Initial response rate was 6 of 85 teachers for a 7% participation rate. In order to increase the participation rate, an incentive was offered. Nulty (2008) asserted there are several ways

to increase the response rate of online surveys. These include: 1) providing a direct URL for the survey via email, 2) providing frequent reminders, 3) persuading respondents that their responses will be used, and 4) provide rewards. Four attempts over a four-week period were made to elicit participation. Only after participants were offered a \$25 Amazon gift card did the response rate climb to 31.7%. There were 584 students assigned to the 27 teachers that consented to be part of the study. A total of 542 had valid fall and spring assessments and were used to analyze the two research questions. Initially there were 584 students assigned to twenty-seven 4th grade classrooms within the data set. Of the initial group of students, 542 had pre- and post- assessments within the school district testing windows. A valid pre-assessment was one that took place within the school district's fall testing window from the start of school until September 30th. The median pre-assessment date was September 2nd. Students that were not in a participating classroom during this timeframe was coded as having "No pre-assessment". Twenty-nine students did not have a pre-assessment within the testing window. A valid post-assessment was one that took place within the school district's spring testing window from April 20th until May 20th. The median pre-assessment date was April 27th. Students that were not in a participating classroom during this timeframe were coded as having "No post-assessment". Eleven students did not have a post-assessment within the testing window. Two students were not in a participating classroom during both the fall and spring testing window.

Instrumentation

Predictor/Independent Variables

Predictor or independent variables were separated into student level inputs and teacher level inputs. Student data were provided by the school district mathematics supervisor. Student ethnicity and sex were derived directly from student enrollment data. Student level independent variables are ethnicity, student educational category, and sex (see Appendix B).

Teacher data were collected via an online survey (see Appendix A) that was emailed to teachers following email confirmation of consent to participate in the research study. The primary purpose of the survey was to determine how teachers grouped students within their mathematics class.

Outcome/Dependent Variable

The dependent variable was the measure of change between the fall 4th grade STAR mathematics assessment and the spring assessment. Two measures of student mathematics achievement were used as dependent variables; Scaled Score (SS) and Grade Equivalent (GE).

The STAR mathematics assessment is an adaptive computer assessment that allows for multiple different test forms, depending on how the student is interacting with the assessment. The STAR assessment provides a conversion from test scores into scale scores by first “estimating each student’s location on the Rasch ability scale, based on the difficulty of the item and by the pattern of right and wrong answers. The assessment then uses a linear transformation to make all scores positive integers ranging from 0 to 1400.” (Learning, 2015). Grade Equivalent is a measure of what a normal grade placement is of

a student of which a particular score is typical. Individual months are in tenths, ranging from 0.0 to 12.9+ (Learning, 2015).

The internal consistency of an instrument is a measure of how well it measures a single construct (Muijs, 2010). Wells and Wollack (2003) state that “professionally developed high-stakes standardized tests should have an internal consistency coefficient of at least .90.” The STAR mathematics assessment has an internal consistency for 4th grade of 0.92 (n=1,306,386) (Learning, 2014). An instrument’s test-retest reliability is its ability to measure the same construct at different times. A test-retest reliability coefficient of $>.80$ is desirable for test used to high-stakes decisions (Muijs, 2010). The STAR mathematics assessment has a retest reliability of 0.83 (n=5,000) (Learning, 2014).

Data Collection and Analyses

Data Collection

Following district approval, expedited IRB approval was granted under #108-SB17-032. The initial recruitment email script (see Appendix C) was modified to include the opportunity of participants to win one of two \$25 Amazon gift cards (see Appendix D). All participants responded within a four-week period (see Table 2: Respondent Rate).

Table 2: Respondent Rate

1 st Week	2 nd Week	3 rd Week	4 th Week	Total
13 respondents	12 respondents	0 respondents	2 respondents	27 respondents

Student assessment data were requested and supplied from the school district at two points within the study. The first was following the completion of the teacher recruitment phase. The district provided fall STAR assessment scores, sex, and ethnicity

data for all students assigned to the participating teachers. Spring STAR assessment scores were provided by the district in May.

Data Analysis

All data analyses were conducted using either SSPS Version 24 (IBM, 2016) and/or Microsoft Excel for Mac Version 15 (2016).

The student data set was provided by the school district. Student mathematics achievement data, educational category, ethnicity, and classroom assignment was provided as a spreadsheet extract from the STAR district level reporting tool. This extract was emailed from the school district as a separate file per each classroom, twenty-seven separate files. The school district also provided a separate roster of student names assigned by teacher. This roster was used to verify that each student was correctly assigned to the proper classroom teacher on the STAR extract. All student data files were transmitted via school district email to researcher's university email account that is password protected. Data cleansing and analysis was conducted on the researcher's password protected computer and stored on researcher's password protected cloud storage. Teacher survey results were extracted directly from Qualtrics (2017). Teacher survey extract was stored on researcher's password protected cloud storage that is provide and maintained by the university.

Student data extracts were first merged into a single spreadsheet file. The teacher indicated mathematics grouping level was then added to the consolidated student spreadsheet as an additional column. Student names and classroom assignments within the consolidated spreadsheet file were compared to classroom rosters to ensure accuracy of the spreadsheet file. At this point, the data were reviewed to determine which students

had valid pre- and post-assessments. Students entries were coded as either having a valid pre- & post-assessment, no pre-assessment, no post-assessment, missing both. The goal was to eliminate students from the analysis that were not assigned to the classroom at both the beginning of the school year when the fall assessment was given and when the spring assessment was given.

At this point, the data were scrubbed to determine whether there was any missingness that needed to be addressed. The first univariate test determined by means of a frequency table (see Table 3: Valid Assessments) that 42 students were missing either a fall assessment, spring assessment, or both. Only students with both pre- and post-assessments were used for further analysis. This gave a total sample size of 542 students.

Table 3: Valid Assessments

	Frequency	Percent
Pre and Post Assessment	542	92.8
No Pre Assessment	29	5.0
No Post Assessment	11	1.9
No Pre and Post Assessment	2	.3
Total	584	100.0

Based on student enrollment data, five ethnic categories were represented within the sample population of this study. The American Indian or Alaskan Native ethnic group was removed from the analysis data set due to their low representation (.9%) of the total sample population and the high variance (standard deviation and standard error of mean) of its five members in regards to the change of scaled score (change_SS) (see Table 4: Change_SS vs Ethnicity).

Table 4: Change_SS vs Ethnicity

Ethnicity	Mean	N	Std. Deviation	Std. Error of Mean
Asian or Pacific Islander	93.08	26	50.462	9.896
Black	62.50	20	65.102	14.557
Hispanic	51.48	75	62.709	7.241
White	85.19	416	61.457	3.013
American Indian or Alaskan Native	115.40	5	84.571	37.821
Total	80.35	542	62.602	2.689

The skewness and kurtosis of both the change in scale score (change_SS) and change in grade equivalent (change_GE) were used as a check of the normality of the distribution (see Table 5: Skewness & Kurtosis). A z-score for the skewness of change_GE yields a value of 13.42 suggesting the scores are skewed to the right. A z-score for the kurtosis of change_GE yields a value of 13.09 suggesting the scores are rather “peaked” and not distributed normally around the mean. Both of these results suggest that change_GE is not normally distributed. The change_SS has both skewness and kurtosis results that suggests those scores are normally distributed.

Table 5: Skewness & Kurtosis

		change_SS	change_GE
N	Valid	537	537
	Missing	0	0
Skewness		.019	1.410
Std. Error of Skewness		.105	.105
Kurtosis		.175	2.749
Std. Error of Kurtosis		.210	.210

Kolmogorov-Smirnov (K-S) and Shapiro-Wilk (S-W) tests were then used to test for normality of distribution for both change_SS and change_GE (see Table 6: Test of Normality – Sex, Table 7: Test of Normality – Ethnicity, Table 8: Test of Normality – Grouping, and Table 9: Test of Normality – Grouping Level). This was used to confirm the previous measures of skewness and kurtosis (Rose, Spinks, & Canhoto, 2015). These tests were accomplished for the dependent variables of change_SS and change_GE across the independent variables of sex, ethnicity, whether or not students were grouped, and grouping level. These tests confirm that change_SS scores are normally distributed across all levels of the independent variables. The distributions of change_GE scores are not normally distributed across the independent variables.

Table 6: Test of Normality - Sex

Tests of Normality							
	Sex	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
change_SS	Female	.054	289	.038	.991	289	.068
	Male	.030	248	.200*	.993	248	.355
change_GE	Female	.130	289	.000	.906	289	.000
	Male	.135	248	.000	.887	248	.000
*. This is a lower bound of the true significance.							
a. Lilliefors Significance Correction							

Table 7: Test of Normality - Ethnicity

Tests of Normality							
	Ethnicity	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
change_SS	Asian or Pacific Islander	.142	26	.188	.953	26	.275
	Black	.115	20	.200*	.947	20	.320
	Hispanic	.070	75	.200*	.986	75	.571
	White	.032	416	.200*	.996	416	.349
change_GE	Asian or Pacific Islander	.167	26	.062	.894	26	.012
	Black	.240	20	.004	.853	20	.006
	Hispanic	.122	75	.008	.968	75	.052
	White	.124	416	.000	.892	416	.000
*. This is a lower bound of the true significance.							
a. Lilliefors Significance Correction							

Table 8: Test of Normality - Grouping

Tests of Normality							
	class_group	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
change_SS	Not Grouped	.050	365	.028	.994	365	.134
	Yes Grouped	.047	172	.200*	.995	172	.872
change_GE	Not Grouped	.120	365	.000	.891	365	.000
	Yes Grouped	.139	172	.000	.900	172	.000
*. This is a lower bound of the true significance.							
a. Lilliefors Significance Correction							

Table 9: Test of Normality - Grouping Level

Tests of Normality							
	teach_grouping	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
change_SS	mixed	.050	365	.028	.994	365	.134
	low	.094	50	.200*	.988	50	.887
	medium	.044	79	.200*	.988	79	.701
	high	.167	43	.004	.960	43	.133
change_GE	mixed	.120	365	.000	.891	365	.000
	low	.167	50	.001	.822	50	.000
	medium	.132	79	.002	.843	79	.000
	high	.092	43	.200*	.959	43	.123
*. This is a lower bound of the true significance.							
a. Lilliefors Significance Correction							

The first research question, “Does ability grouping affect mathematics achievement of 4th grade students?”, was analyzed two different ways. An independent samples *t*-test was used to explore whether mathematics achievement was dependent on whether or not students were in ability groups. An analysis of variance (ANOVA) was used to explore whether mathematics achievement was dependent on the level of mathematics group. According to Salkind (2010), inferential statistical methods, such as chi-square test, paired-samples and independent *t* tests, and analysis of variance (ANOVA), are appropriate within a casual-comparative research study when the “researcher hopes to demonstrate that a relationship exist between the independent and dependent variables.”

There are three assumptions that must be met in order use an independent samples *t* test and ANOVAs. The first is an assumption of independence that states that the two groups under analysis are independent of one another. In the case of this study, students

were either ability grouped or they were not. There was not an instance where students were in both conditions. The second assumption is that of normality. This assumes the dependent variable is normally distributed within each of the two groups. This was tested by use of the Shapiro-Wilks test. The final assumption is that homogeneity of variance. This assumes the variances of the dependent variable within the each of the two groups are equal. This was tested by use of Levene's *F* Test for Equality.

The second research question, "Does ability grouping affect the mathematics achievement of 4th grade students based on the sex and ethnicity of students?", was analyzed using ANOVAs and a linear regression. Analysis of variance (ANOVA) allows comparison of a continuous dependent variable and several groups (Muijs, 2010). Main effects and interactions were analyzed on the scale score change (change_SS) dependent variable and the independent variables of grouping level, sex, and ethnicity.

A multiple regression was conducted to assist in measuring the association between the independent variables of sex, ethnicity, and grouping level and the dependent variable, change in scale score (change_SS). This will help in the construction of a linear equation that will predict values of the dependent variable based on the independent variables (Gray & Kinnear, 2012). There are four assumptions that must be met in order to use a linear regression. The first is the linearity of residuals. The second is independence of residuals. The third is that the residuals are normally distributed. The final assumption is that the residuals have equal variance.

The next chapter will use the preceding methodology to determine the relationships between mathematics achievement and the independent variables of sex, ethnicity, and ability grouping.

CHAPTER 4: RESULTS

Overview of the Study

A causal-comparative study was used to study the effects that grouping, either homogenously or heterogeneously, had on the mathematics achievement of 4th grade students. Student achievement was measured from the beginning of 4th grade till the 4th quarter of the school year based on the STAR (Learning, 2014, 2015) standardized mathematics assessment.

There were two guiding research questions and associated hypothesis for this study.

Research Question 1 - Does ability grouping affect mathematics achievement of 4th grade students?

H₀: Ability grouping does not have an affect on mathematics achievement of 4th grade students.

H₁: Ability grouping does have an affect on mathematics achievement of 4th grade students.

Research Question 2 - Does ability grouping affect the mathematics achievement of 4th grade students differently based on sex and ethnicity of students?

H₀: Ability grouping by mathematics achievement of 4th grade students does not affect students differently by sex and ethnicity.

H₁: Ability grouping by mathematics achievement of 4th grade students affects students differently by sex and ethnicity.

Data Analyses

Data Cleansing

All analyses were conducted using IBM SPSS Version 24 for Macintosh (IBM, 2016). Post hoc power analyses for independent sample t-test were conducted using G*Power 3.1 (Faul et al., 2009).

Post hoc power analysis using G*Power 3.1 (Faul et al., 2009) was conducted for the two-tailed independent samples t-test. Power was calculated to be 0.7793600. This equates an 78% probability of detecting a true difference between females and male students (see Table 10: T-test Post hoc Power Computation).

Table 10: T-Test Post hoc Power Computation

t tests - Means: Difference between two independent means (two groups)			
Analysis:	Post hoc: Compute achieved power		
Input:	Tail(s)	=	Two
	Effect size d	=	0.2357226
	α err prob	=	0.05
	Sample size group 1	=	293
	Sample size group 2	=	249
Output:	Noncentrality parameter δ	=	2.7348598
	Critical t	=	1.9643668
	Df	=	540
	Power (1- β err prob)	=	0.77936

Student data, including sex, ethnicity, and mathematics achievement scores, were provided by the participating school district. Initially there were 584 students assigned to twenty-seven 4th grade classrooms within the data set. Of the initial group of students, 542 had pre- and post- assessments within the school district testing windows (see Table 11: Valid Assessments). A valid pre-assessment was one that took place within the school

district's fall testing window from the start of school until September 30th. The median pre-assessment date was September 2nd. Students that were not in a participating classroom during this timeframe was coded as having "No pre-assessment". Twenty-nine students did not have a pre-assessment within the testing window. A valid post-assessment was one that took place within the school district's spring testing window from April 20th until May 20th. The median pre-assessment date was April 27th. Students that were not in a participating classroom during this timeframe was coded as having "No post-assessment". Eleven students did not have a post-assessment within the testing window. Two students were not in a participating classroom during both the fall and spring testing window.

Table 11: Valid Assessments

	Frequency	Percent
Pre and Post Assessment	542	92.8
No Pre Assessment	29	5.0
No Post Assessment	11	1.9
No Pre and Post Assessment	2	.3
Total	584	100.0

Of the 542 remaining students with assessments within the testing window, 54% were female and 46% male (see Table 12: Sex).

Table 12: Sex

	Frequency	Percent
Female	293	54.1
Male	249	45.9
Total	542	100.0

Based on student enrollment data, five ethnic categories were represented within the study participants. The American Indian or Alaskan Native ethnic group was removed from the analysis data set due to their low representation (.9%) of the total sample and the high variance (standard deviation and standard error of mean) of its five members in regards to the change of scaled score (change_SS) (see Table 13: Change_SS vs Ethnicity). This brings the total student participants to 537.

Table 13: Change_SS vs Ethnicity

Ethnicity	Mean	N	Std. Deviation	Std. Error of Mean
Asian or Pacific Islander	93.08	26	50.462	9.896
Black	62.50	20	65.102	14.557
Hispanic	51.48	75	62.709	7.241
White	85.19	416	61.457	3.013
American Indian or Alaskan Native	115.40	5	84.571	37.821
Total	80.35	542	62.602	2.689

Research Questions #1 - Grouping

The first research question addressed by this study is, “Does ability grouping affect mathematics achievement of 4th grade students?”

The hypotheses of this the first research question are:

H₀: Ability grouping does not have an affect on mathematics achievement of 4th grade students.

H₁: Ability grouping does have an affect on mathematics achievement of 4th grade students.

Grouping

The first test conducted was to determine whether the means (see Table 5: *Achievement vs. Grouping*) of both the change in scale score (change_SS) and the change in grade equivalency (change_GE) was significant for groups that were homogeneously grouped (yes grouped) or heterogeneously grouped (not grouped).

An independent samples *t*-test was conducted to test the hypothesis of research question 1. There are three assumptions that must be met in order to use a *t*-test or an ANOVA. The first is an assumption of independence that states that the two groups under analysis are independent of one another. In the case of this study, students were either ability grouped or they were not. There was not an instance where students were in both conditions. This assumption of independence is valid for both the dependent variables of change in scale score (change_SS) and change in grade equivalency (change_GE).

The second assumption is that of normality. This assumes that the dependent variable is normally distributed within each of the two groups. This assumption was first tested for the dependent variable of change_SS. The assumption of normality for the “not grouped” students was tested by the Shapiro-Wilks test for normality (SW = .994, df = 365, $p = .134$) and skewness (.098) and kurtosis (.240). These tests suggested that normality was a reasonable assumption. The assumption of normality for the “grouped” students was tested by the Shapiro-Wilks test for normality (SW = .995, df = 172, $p = .872$), skewness (-.170) and kurtosis (.036). These statistics suggested that normality was a reasonable assumption.

The assumption of normality for the dependent variable of change_GE highlighted the non-normal distribution of these scores. The assumption of normality for

the “not grouped” students was tested by the Shapiro-Wilks test for normality ($SW = .891$, $df = 365$, $p < .001$), skewness (1.555) and kurtosis (3.880). These tests suggested that these scores were not normally distributed. A nonparametric test was used to analyze change_GE given its non-normal distribution. The assumption of normality for the “grouped” students was tested by the Shapiro-Wilks test for normality ($SW = .900$, $df = 172$, $p < .001$), skewness (1.091) and kurtosis (1.206). These statistics suggested that these scores were also not normally distributed.

The final assumption is that of homogeneity of variance. Levene’s test for equality of variances for change_SS indicated equality of variances ($F = .239$, $p = .625$). The Levene’s test of variance for change_GE indicated unequal variances ($F = 12.461$, $p < .001$).

The means for the change in scaled score (change_SS) for the grouped students were higher ($M = 81.86$, $SD = 59.830$) than for the not grouped students ($M = 79.15$, $SD = 63.601$) (see Table 14: Achievement vs. Grouping). The t -test indicates that there is not a statistical significance ($t(535) = .469$, $p = .639$, $d = 0.44$) between the change in scaled score between students that are grouped and not grouped.

Table 14: Achievement vs. Grouping

Group Statistics					
	class_group		Mean	Std. Deviation	Std. Error Mean
change_SS	Not Grouped	65	79.15	63.601	3.329
	Yes Grouped	72	81.86	59.830	4.562
change_GE	Not Grouped	65	1.4447	1.42921	.07481
	Yes Grouped	72	1.8267	1.85248	.14125

The change in grade equivalency (change_GE) did not pass the assumption of normality and the assumption of homogeneity of variance. In this case, a nonparametric

test that requires fewer assumptions of equality is appropriate (Gray & Kinnear, 2012). Two nonparametric alternatives to the independent-samples t-test are the Mann-Whitney U test and the Wilcoxon Rank Sum test. The change in grade equivalency (change_GE) of not grouped students (N=365, M=1.44, SD=1.429) were not significantly different than that of the grouped students (N=172, M=1.85, SD=1.852) (see Table 5). A Mann-Whitney U test showed the difference to be non-significant: $U = 34,563.50$; $p = .058$ (two-tailed).

The null hypothesis, “Ability grouping does not have an affect on mathematics achievement of 4th grade students.”, would not be rejected if one was to just look at the research question from whether or not grouping students effect math achievement.

An analysis of variance (ANOVA) was conducted to determine whether there was a significant difference in growth between each of the four groups (below grade level = low, at grade level = medium, above grade level=high, heterogeneous grouped = mixed). Students were only in one of the four ability groups and were not in multiple grouping levels. This is key for the assumption of independence to be met. The low group had fifty students within it, the medium group had seventy-nine students, the high group had forty-three students, and the mixed group had three hundred and sixty-five students (see Table 6). The composition of sex and ethnicity by type of group is found in Table 23. There was no significant difference between the representation of females and males within each of the four ability groups (chi square = .822, $df = 3$, $p = .844$). There was a significant difference between the representation of ethnicity within the four ability groups (chi square = 41.395, $df = 9$, $p < 000$). The Asian/Pacific Islander group was overrepresented in the above grade level group and underrepresented in the below grade

level group. The Black group had no students within the below grade level group. Finally, the Hispanic group was not proportionally represented within the above grade level group. This is not surprising, as assignment to ability groups would not take place based on ethnic representations. The factors used to assign students to ability groups was not addressed within this study.

The assumption of normality for the dependent variable of change_SS for the “below grade level” group was tested by the Shapiro-Wilks test for normality (SW = .988, df = 50, p = .887), skewness (.302) and kurtosis (.003). These tests suggested that normality was a reasonable assumption. Assumption of normality for the “grade level” group suggested that these scores were normally distributed (SW = .988, df = 79, p = .701, skewness (-.244) and kurtosis (-.057)). Assumption of normality for the “above grade level” group suggested that these scores were normally distributed (SW = .960, df = 43, p = .133, skewness (-.290) and kurtosis (-.717)). Finally, the assumption of normality for the “heterogeneous grouped” students suggested that these scores were normally distributed (SW = .994, df = 365, p = .134, skewness (.098) and kurtosis (.240)).

The assumption of normality for the dependent variable of change_GE for the “below grade level” group was tested by the Shapiro-Wilks test for normality (SW = .822, df = 50, p < .001), skewness (1.865) and kurtosis (4.149). These tests suggested that these scores were not distributed normally. Assumption of normality for the “grade level” group suggested that these scores were not normally distributed (SW = .843, df = 79, p < .001, skewness (2.038) and kurtosis (6.643)). Assumption of normality for the “above grade level” group suggested that these scores were normally distributed (SW = .959, df

= 43, $p = .123$, skewness (-.314) and kurtosis (.018)). Finally, the assumption of normality for the “heterogeneous grouped” students suggested that these scores were not normally distributed (SW = .891, $df = 365$, $p < .001$, skewness (1.555) and kurtosis (3.880)).

The final assumption is that of homogeneity of variance. Levene’s test for equality of variances for change_SS indicated equality of variances ($F = .496$, $p = .685$). The Levene’s test of variance for change_GE indicated unequal variances ($F = 8.024$, $p < .001$). An analysis of variance was only conducted for the dependent variable of change_SS and not of change_GE. This is due to fact that the means for change_GE violated the assumptions of normality and equality of variance.

An analysis of variance showed a main effect of grouping on change in scaled score (change_SS), $F(3,533) = 4.407$, $p = .004$, $\eta_p^2 = .024$. Post-hoc analyses using Tukey’s HSD indicated that the change_SS had a statistically lower mean ($p = .005$) for the medium grouped students ($N=79$, $M=63.73$, $SD=60.517$) than the low grouped students ($N=50$, $M=101.12$, $SD=56.033$) (see Table 15: Change_SS vs Levels of Grouping). Observed power for the analysis of variance was .874. The model weakly accounts for the variance in the change in a student’s scaled score based solely on how they are grouped ($R^2 = 0.024$).

The null hypothesis, “Ability grouping does not have an affect on mathematics achievement of 4th grade students.”, should not be rejected based on the preceding analysis.

Table 15: Change_SS vs Levels of Grouping

Dependent Variable: change_SS			
teach_grouping	Mean	Std. Deviation	N
Mixed	79.15	63.601	65
Below Grade Level	101.12	56.033	0
At Grade Level	63.73	60.517	9
Above Grade Level	92.77	54.112	3
Total	80.02	62.375	37

Research Questions #2 – Ethnicity and Sex

The second research question addressed by this study is, “Does ability grouping affect differently mathematics achievement of 4th grade students based on the ethnicity and sex of students?”

The hypotheses of this second research question are:

H₀: Ability grouping by mathematics achievement of 4th grade students does not affect students differently by sex and ethnicity.

H₁: Ability grouping by mathematics achievement of 4th grade students affects students differently by sex and ethnicity.

Ethnicity

Ethnicity was analyzed first using an analysis of variance to look for significant differences between ethnic groups based on the means of changed scaled score (change_SS) (see Table 16: Ethnicity vs Change_SS).

Students were only in one of the four ethnic groups and were not in multiple groups. This limitation is inherited from the way students’ ethnicity is coded within STAR. This is key in order for the assumption of independence to be met.

The assumption of normality for the dependent variable of change_SS for the “Asian or Pacific Islander” group was tested by the Shapiro-Wilks test for normality (SW = .953, df = 26, p = .275), skewness (.273) and kurtosis (-.951). These tests suggested that normality was a reasonable assumption. Assumption of normality for the “Black” group suggested that these scores were normally distributed (SW = .947, df = 20, p = .320, skewness (.529) and kurtosis (-.617)). Assumption of normality for the “Hispanic” group suggested that these scores were normally distributed (SW = .986, df = 75, p = .571, skewness (-.013) and kurtosis (.956)). Finally, the assumption of normality for the “White” students suggested that these scores were normally distributed (SW = .996, df = 416, p = .349, skewness (.037) and kurtosis (.124)).

The assumption of normality for the dependent variable of change_GE for the “Asian or Pacific Islander” group was tested by the Shapiro-Wilks test for normality (SW = .894, df = 26, p = .012), skewness (1.207) and kurtosis (1.408). These statistics suggested that these scores were not distributed normally. Assumption of normality for the “Black” group suggested that these scores were not normally distributed (SW = .853, df = 20, p = .006, skewness (.749) and kurtosis (-.999)). Assumption of normality for the “Hispanic” group suggested that these scores were normally distributed (SW = .968, df = 75, p = .052, skewness (.487) and kurtosis (1.449)). Finally, the assumption of normality for the “White” students suggested that these scores were not normally distributed (SW = .892, df = 416, p < .001, skewness (1.351) and kurtosis (2.424)).

The final assumption is that of homogeneity of variance. Levene’s test for equality of variances for change_SS indicated equality of variances for ethnicity (F = .342, p = .795). The Levene’s test of variance for change_GE indicated unequal variances

for ethnicity ($F= 5.903$, $p = .001$). An analysis of variance was only conducted for the dependent variable of change_SS and not of change_GE. This is due to fact that the means for change_GE violated the assumptions of normality and equality of variance.

Analysis of variance showed a main effect on the change in scaled score for ethnicity, $F(3,533) = 7.345$, $p < .001$, $\eta p^2 = .040$. Post- hoc analyses using Tukey's HSD indicated that mathematics achievement was lower for Hispanics than Asian or Pacific Islander students ($p=.016$) and lower for Hispanics than White students ($p<.001$). Observed power for the analysis of variance was .985. The model weakly accounts for the variance in the change in a student's scaled score based on their ethnicity (R Squared = 0.040).

Table 16: Ethnicity vs Change_SS

Dependent Variable: change_SS			
Ethnicity	Mean	Std. Deviation	N
Asian or Pacific Islander	93.08	50.462	26
Black	62.50	65.102	20
Hispanic	51.48	62.709	75
White	85.19	61.457	416
Total	80.02	62.375	537

The dependent variable of change_SS was looked at in terms of ethnicity and the four ability groups the students were placed in. There were two groups of students that had less than 5 participants per cell (see Table 17: Ethnicity and Grouping vs. Change_SS). These groups were Asian/Pacific Islander in the medium group ($N=1$) and Black in the high group ($N=1$). A sample size this small makes it problematic to state this group represents the population. For this reason, further analysis was not performed based on ethnicity and grouping level. Instead, ethnicity compared to whether or not

students were grouped (see Table 18: Ethnicity and Grouping Level vs Change_SS). An analysis of variance showed a main effect on the change in scaled score for ethnicity, $F(7,537) = 3.831, p < .001, \eta^2 = .048$. The interaction between ethnicity and whether or not students were grouped were found to be non-significant ($F(3, 537) = 1.565, p = .197, \eta^2 = .009$). Observed power for the analysis of variance was .981. The model weakly accounts for the variance in the change in a student's scaled score based on their ethnicity (R Squared = 0.048).

Table 17: Ethnicity and Grouping vs Change_SS

Dependent Variable: change_SS				
Ethnicity	teach_grouping	Mean	Std. Deviation	N
Asian or Pacific Islander	mixed	76.50	47.531	12
	below grade level	115.80	44.584	5
	at grade level	145.00		1
	above grade level	97.25	56.244	8
	Total	93.08	50.462	26
Black	mixed	51.08	69.019	13
	at grade level	77.67	58.277	6
	above grade level	120.00		1
	Total	62.50	65.102	20
Hispanic	mixed	49.17	65.990	64
	below grade level	49.25	26.094	4
	at grade level	73.86	42.314	7
	Total	51.48	62.709	75
White	mixed	87.54	61.164	276
	below grade level	104.39	57.353	41
	at grade level	60.11	62.324	65
	above grade level	90.91	54.998	34
	Total	85.19	61.457	416
Total	mixed	79.15	63.601	365
	below grade level	101.12	56.033	50
	at grade level	63.73	60.517	79
	above grade level	92.77	54.112	43
	Total	80.02	62.375	537

Table 18: Ethnicity and Grouped vs Change_SS

Dependent Variable: change_SS				
Ethnicity	class_group	Mean	Std. Deviation	N
Asian or Pacific Islander	Not Grouped	76.50	47.531	12
	Yes Grouped	107.29	50.142	14
	Total	93.08	50.462	26
Black	Not Grouped	51.08	69.019	13
	Yes Grouped	83.71	55.554	7
	Total	62.50	65.102	20
Hispanic	Not Grouped	49.17	65.990	64
	Yes Grouped	64.91	37.851	11
	Total	51.48	62.709	75
White	Not Grouped	87.54	61.164	276
	Yes Grouped	80.56	61.988	140
	Total	85.19	61.457	416
Total	Not Grouped	79.15	63.601	365
	Yes Grouped	81.86	59.830	172
	Total	80.02	62.375	537

Sex

Sex was analyzed first using an *f*-test to look for significant differences between female and male students based on the means of changed scaled score (change_SS) (see Table 19: Sex vs Change_SS).

Students were identified as being either male or female, not both. Thus, the assumption of independence is met.

The assumption of normality for the dependent variable of change_SS for the “female” group was tested by the Shapiro-Wilks test for normality (SW = .991, df = 289, $p = .068$), skewness (.258) and kurtosis (.183). These tests suggested that normality was a reasonable assumption. Assumption of normality for the “male” group suggested that these scores were normally distributed (SW = .993, df = 248, $p = .355$, skewness (-.257) and kurtosis (.456)).

The assumption of normality for the dependent variable of change_GE for the “female” group was tested by the Shapiro-Wilks test for normality (SW = .906, df = 289, $p < .001$), skewness (1.304) and kurtosis (3.259). These tests suggested that these scores were not distributed normally. Assumption of normality for the “male” group suggested that these scores were not normally distributed (SW = .887, df = 248, $p < .001$, skewness (1.357) and kurtosis (1.926)).

The final assumption is that of homogeneity of variance. Levene’s test for equality of variances for change_SS indicated equality of variances for sex ($F = .267$, $p = .605$). The Levene’s test of variance for change_GE indicated unequal variances for sex ($F = 7.142$, $p = .008$). An analysis of variance was only conducted for the dependent variable of change_SS and not of change_GE. This is due to fact that the means for change_GE violated the assumptions of normality and equality of variance.

Analysis of variance showed a main effect on the change in scaled score for sex, $F(1,537) = 7.549$, $p = .006$, $\eta_p^2 = .014$. Observed power for the analysis of variance was .783. The model weakly accounts for the variance in the change in a student’s scaled score based on their sex (R Squared = 0.014). The difference between the mathematics achievement of female and male students within this sample was statistically significant.

Table 19: Sex vs Change_SS

Dependent Variable: change_SS			
Sex	Mean	Std. Deviation	N
Female	73.21	61.339	289
Male	87.96	62.757	248
Total	80.02	62.375	537

A multi-factor ANOVA was performed to determine the effect that both sex and

grouping level had on the change in scaled score (change_SS) (see Table 20: Sex and Grouping vs Change_SS). A 4x2 ANOVA with sex (female and male) and grouping level (mixed, low, medium, high) as between-subjects factors revealed a main effect, $F(7, 537) = 3.304$, $p = .002$, $\eta_p^2 = .042$, of sex, $F(1, 537) = 1.449$, $p = .229$, $\eta_p^2 = .003$, and grouping, $F(3, 537) = 4.267$, $p = .005$, $\eta_p^2 = .024$. The interaction between these main effects were not statistically significant, $F(3, 537) = .917$, $p = .433$, $\eta_p^2 = .005$. Observed power for the analysis of variance was .959. The model weakly accounts for the variance in the change in a student's scaled score based on their sex and grouping level (R Squared = 0.042).

Table 20: Sex and Grouping vs Change_SS

Dependent Variable: change_SS				
Sex	teach_grouping	Mean	Std. Deviation	N
Female	mixed	71.31	59.803	198
	below grade level	108.04	60.267	24
	at grade level	56.20	61.750	44
	above grade level	85.74	62.296	23
	Total	73.21	61.339	289
Male	mixed	88.45	66.825	167
	below grade level	94.73	52.191	26
	at grade level	73.20	58.428	35
	above grade level	100.85	43.018	20
	Total	87.96	62.757	248
Total	mixed	79.15	63.601	365
	below grade level	101.12	56.033	50
	at grade level	63.73	60.517	79
	above grade level	92.77	54.112	43
	Total	80.02	62.375	537

Sex and Ethnicity

A multi-factor ANOVA was performed to determine the effect that both sex and

ethnicity had on the change in scaled score (change_SS) (see Table 21: Sex and Ethnicity vs Change_SS). A 2x4 ANOVA with sex (female and male) and ethnicity (Asian or Pacific Islander, Black, Hispanic, White) as between-subjects factors revealed a main effects of sex, $F(1, 537) = 1.037$, $p = .309$, $\eta_p^2 = .002$, and ethnicity, $F(3, 537) = 7.201$, $p < .001$, $\eta_p^2 = .039$. Levene's test indicated equal variances ($F = .450$, $p = .870$). The interaction between these main effects were not statistically significant, $F(3, 537) = 1.095$, $p = .351$, $\eta_p^2 = .006$. Observed power for the analysis of variance was .996. The model weakly accounts for the variance in the change in a student's scaled score based on their sex and ethnicity (R Squared = 0.059).

Table 21: Sex and Ethnicity vs Change_SS

Dependent Variable: change_SS				
Sex	Ethnicity	Mean	Std. Deviation	N
Female	Asian or Pacific Islander	102.69	58.967	13
	Black	51.00	57.918	14
	Hispanic	48.90	59.304	39
	White	77.14	60.772	223
	Total	73.21	61.339	289
Male	Asian or Pacific Islander	83.46	40.344	13
	Black	89.33	78.419	6
	Hispanic	54.28	66.936	36
	White	94.50	61.081	193
	Total	87.96	62.757	248
Total	Asian or Pacific Islander	93.08	50.462	26
	Black	62.50	65.102	20
	Hispanic	51.48	62.709	75
	White	85.19	61.457	416
	Total	80.02	62.375	537

Sex, Ethnicity, and Grouping

The dependent variable of change_SS was looked at in terms of sex, ethnicity,

and whether or not students were grouped. There were three groups of students that less than 5 participants per cell (see Table 22: Sex, Ethnicity, and Grouped vs. Change_SS). Once again, a sample size this small makes it problematic to state this group represents the population. These groups were 1) female – Hispanic – “Yes Grouped” (N = 4), 2) male – Black – “Yes Grouped” (N = 2), and 3) male – Black – Not Grouped” (N = 4). For this reason, further analysis was not performed based on ethnicity and grouping level.

Table 22: Sex, Ethnicity, and Grouped vs Change_SS

Dependent Variable: change_SS					
Sex	Ethnicity	class_group	Mean	Std. Deviation	N
Female	Asian or Pacific Islander	Not Grouped	87.50	59.989	6
		Yes Grouped	115.71	59.385	7
		Total	102.69	58.967	13
	Black	Not Grouped	29.11	48.735	9
		Yes Grouped	90.40	55.976	5
		Total	51.00	57.918	14
	Hispanic	Not Grouped	49.00	61.804	35
		Yes Grouped	48.00	35.449	4
		Total	48.90	59.304	39
	White	Not Grouped	78.50	58.029	148
		Yes Grouped	74.45	66.178	75
		Total	77.14	60.772	223
	Total	Not Grouped	71.31	59.803	198
		Yes Grouped	77.34	64.702	91
		Total	73.21	61.339	289
Male	Asian or Pacific Islander	Not Grouped	65.50	32.880	6
		Yes Grouped	98.86	41.894	7
		Total	83.46	40.344	13
	Black	Not Grouped	100.50	89.534	4
		Yes Grouped	67.00	72.125	2
		Total	89.33	78.419	6
	Hispanic	Not Grouped	49.38	71.832	29
		Yes Grouped	74.57	38.209	7
		Total	54.28	66.936	36

	White	Not Grouped	98.00	63.225	128
		Yes Grouped	87.60	56.458	65
		Total	94.50	61.081	193
	Total	Not Grouped	88.45	66.825	167
		Yes Grouped	86.94	53.780	81
		Total	87.96	62.757	248
Total	Asian or Pacific Islander	Not Grouped	76.50	47.531	12
		Yes Grouped	107.29	50.142	14
		Total	93.08	50.462	26
	Black	Not Grouped	51.08	69.019	13
		Yes Grouped	83.71	55.554	7
		Total	62.50	65.102	20
	Hispanic	Not Grouped	49.17	65.990	64
		Yes Grouped	64.91	37.851	11
		Total	51.48	62.709	75
	White	Not Grouped	87.54	61.164	276
		Yes Grouped	80.56	61.988	140
		Total	85.19	61.457	416
	Total	Not Grouped	79.15	63.601	365
		Yes Grouped	81.86	59.830	172
		Total	80.02	62.375	537

The dependent variable of change_SS was looked at in terms of sex, ethnicity, and grouping level. There were twelve groups of students that less than 5 participants per cell (see Table 23: Sex, Ethnicity, and Grouping Level vs. Change_SS). For this reason, further analysis was not performed based on ethnicity and grouping level.

Table 23: Sex, Ethnicity, and Grouping Level vs Change_SS

Sex	Ethnicity	teach_grouping	Mean	Std. Deviation	N
Female	Asian or Pacific Islander	mixed	87.50	59.989	6
		below grade level	147.00	28.284	2
		above grade level	103.20	66.368	5
		Total	102.69	58.967	13
	Black	mixed	29.11	48.735	9
		at grade level	83.00	61.747	4

		above grade level	120.00		1
		Total	51.00	57.918	14
	Hispanic	mixed	49.00	61.804	35
		below grade level	51.50	43.134	2
		at grade level	44.50	43.134	2
		Total	48.90	59.304	39
	White	mixed	78.50	58.029	148
		below grade level	109.80	61.296	20
		at grade level	54.00	63.102	38
		above grade level	78.59	63.339	17
		Total	77.14	60.772	223
	Total	mixed	71.31	59.803	198
		below grade level	108.04	60.267	24
		at grade level	56.20	61.750	44
		above grade level	85.74	62.296	23
Total		73.21	61.339	289	
Male	Asian or Pacific Islander	mixed	65.50	32.880	6
		below grade level	95.00	44.193	3
		at grade level	145.00		1
		above grade level	87.33	45.015	3
		Total	83.46	40.344	13
	Black	mixed	100.50	89.534	4
		at grade level	67.00	72.125	2
		Total	89.33	78.419	6
	Hispanic	mixed	49.38	71.832	29
		below grade level	47.00	12.728	2
		at grade level	85.60	40.216	5
		Total	54.28	66.936	36
	White	mixed	98.00	63.225	128
		below grade level	99.24	54.335	21
		at grade level	68.70	61.348	27
		above grade level	103.24	43.634	17
		Total	94.50	61.081	193
	Total	mixed	88.45	66.825	167
		below grade level	94.73	52.191	26
		at grade level	73.20	58.428	35
		above grade level	100.85	43.018	20
		Total	87.96	62.757	248

Total	Asian or Pacific Islander	mixed	76.50	47.531	12
		below grade level	115.80	44.584	5
		at grade level	145.00		1
		above grade level	97.25	56.244	8
		Total	93.08	50.462	26
	Black	mixed	51.08	69.019	13
		at grade level	77.67	58.277	6
		above grade level	120.00		1
		Total	62.50	65.102	20
	Hispanic	mixed	49.17	65.990	64
		below grade level	49.25	26.094	4
		at grade level	73.86	42.314	7
		Total	51.48	62.709	75
	White	mixed	87.54	61.164	276
		below grade level	104.39	57.353	41
		at grade level	60.11	62.324	65
		above grade level	90.91	54.998	34
		Total	85.19	61.457	416
	Total	mixed	79.15	63.601	365
		below grade level	101.12	56.033	50
at grade level		63.73	60.517	79	
above grade level		92.77	54.112	43	
Total		80.02	62.375	537	

Ideally, a multiple linear regression would be used to predict the change in scale score (change_SS) based on a student's sex, ethnicity, and math grouping. This was unable to be accomplished given the lack of representation as students were put in categories based on their sex, ethnicity, and grouping level. There was only sufficient representation to analysis sex and grouping level versus the change_SS using a linear regression model.

There are four assumptions that must be met in order to use a linear regression. The first is the linearity of residuals. The second is independence of residuals. The third

is that the residuals are normally distributed. The final assumption is that the residuals have equal variance.

The assumption of linearity of residuals was checked by examining the scatterplot of residuals across the range of predicated values (see Figure 1: Scatterplot of Residuals).

This plot suggests that the assumption of linearity has been met.

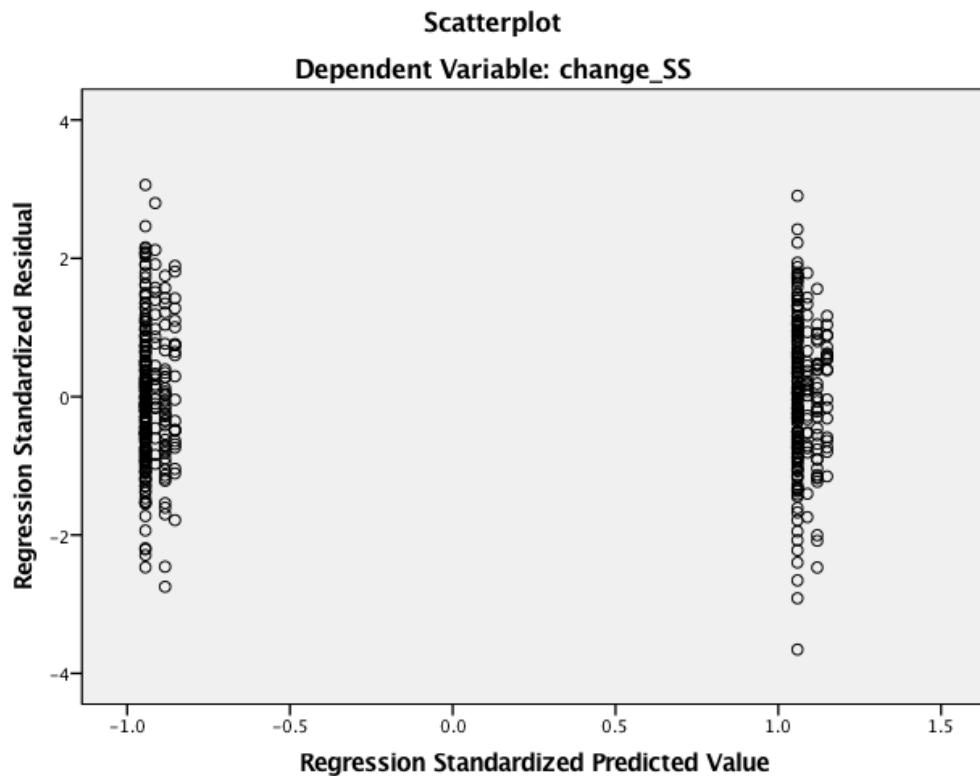


Figure 1: Scatterplot of Residuals

The assumption of independence of residuals has been met by the cross-sectional design of this study. Students were only in one math group at a time and were either male or female.

The assumption of normality of residuals was checked by examining the histogram of the regression standardized residuals (see Figure 2: Histogram of Residuals). Observed points suggest that the assumption of normality of residuals is met.

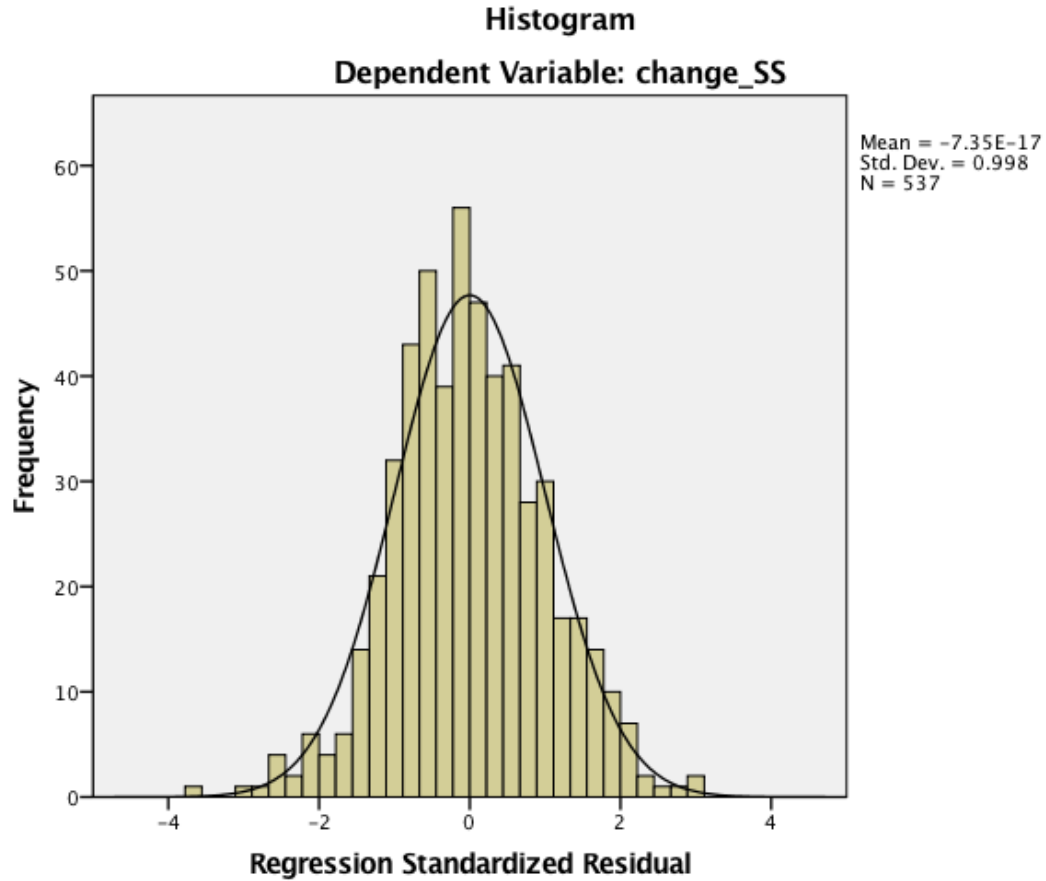


Figure 2: Histogram of Residuals

The final assumption is that of equal variance of residuals. The normal plot of standardized regression residuals suggests that the regression model meets the assumption of “homoscedasticity” or homogeneity of variance (see Figure 3: Plot of Regression Standardized Residuals).

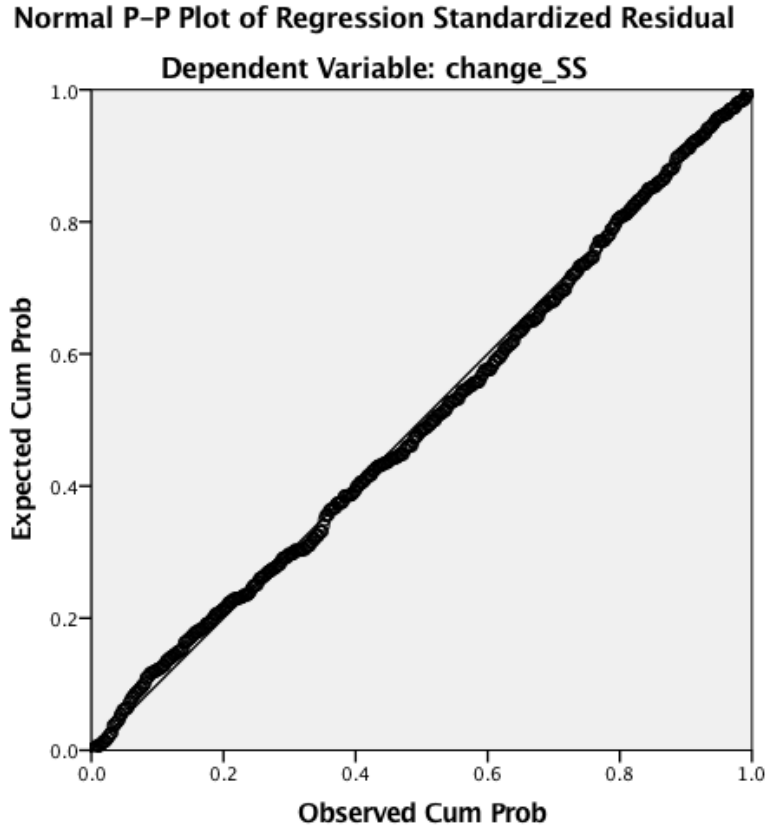


Figure 3: Plot of Regression Standardized Residuals

Overall, the regression model significantly predicted the outcome variable of change in scale score ($F(2,536) = 3.771, p = .024$), with an R^2 of .014. Sex had a significant beta coefficient = .118, $t = 2.745$, and $p = .006$. Grouping had a non-significant beta coefficient = .004, $t = .084$, $p = .933$. Student's predicted change in scale score (change_SS) is equal to the equation $Y^1 = 58.327 + 14.744 (\text{Sex}) + .223 (\text{Group})$, where sex is coded as 1 = Female, 2 = Male, and grouping is coded as 0 = mixed, 1 = low, 2 = medium, and 3 = high. Male students increased their scale score 14.744 points more than females when controlling for grouping. The independent variables do not violate assumption of multicollinearity. Sex and grouping had a correlation coefficient of .001 (see Table 24: Regression Correlations). The collinearity tolerance for sex was 1.000 and grouping was 1.000. All of these figures are close to 1 and are another confirmation that

the assumption of multicollinearity was not violated (Muijs, 2010). The Mahalanobis Distance (M-D) maximum was 6.727 (see Table 25: Residuals Statistics), which exceeds the recommended distance for two independent variables of 13.82. This indicates that there are no possible outliers within the sample. The model was particularly unsuccessful in predicating the change in scale score for two cases, 181 and 370 (see Table 26: Casewise Diagnostics). The Cook's Distance of 0.021 indicate that cases 181 and 370 are not putting an undue stress on the model (see Table 25: Residuals Statistics).

Table 24: Regression Correlations

		change_SS	Sex	teach_grouping
Pearson Correlation	change_SS	1.000	.118	.004
	Sex	.118	1.000	.001
	teach_grouping	.004	.001	1.000
Sig. (1-tailed)	change_SS	.	.003	.465
	Sex	.003		.487
	teach_grouping	.465	.487	
N	change_SS	537	537	537
	Sex	537	537	537
	teach_grouping	537	537	537

Table 25: Residuals Statistics

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	73.07	88.48	80.02	7.361	537
Std. Predicted Value	-.944	1.150	.000	1.000	537
Standard Error of Predicted Value	3.784	7.450	4.542	.939	537
Adjusted Predicted Value	72.09	89.53	80.02	7.369	537
Residual	-226.815	189.929	.000	61.939	537
Std. Residual	-3.655	3.061	.000	.998	537
Stud. Residual	-3.664	3.067	.000	1.001	537
Deleted Residual	-227.901	190.727	.002	62.277	537
Stud. Deleted Residual	-3.707	3.092	.000	1.003	537

Mahal. Distance	.995	6.727	1.996	1.460	537
Cook's Distance	.000	.021	.002	.003	537
Centered Leverage Value	.002	.013	.004	.003	537
a. Dependent Variable: change_SS					

Table 26: Casewise Diagnostics

Case Number	Std. Residual	change_SS	Predicted Value	Residual
181	-3.655	-139	87.82	-226.815
370	3.061	263	73.07	189.929
a. Dependent Variable: change_SS				

The next chapter will detail the interpretation of findings, practical implications for educators, and finally recommendations for further study.

CHAPTER 5: CONCLUSION

Summary

Purpose

The purpose of this quantitative study was to compare the mathematics achievement of 4th grade elementary school students who were grouped in either homogenous or heterogeneous mathematics classes. Student achievement was measured from the beginning of 4th grade till the 4th quarter of the school year based on the STAR standardized mathematics assessment (Learning, 2014). Variations in growth were controlled for by student ethnicity, student sex and with the primary interest on the effects of ability grouping of students on math achievement.

There were two guiding research questions and associated hypothesis for this study.

Research Question 1 - Does ability grouping affect mathematics achievement of 4th grade students?

H₀: Ability grouping does not have an affect on mathematics achievement of 4th grade students.

H₁: Ability grouping does have an affect on mathematics achievement of 4th grade students.

Research Question 2 - Does ability grouping affect the mathematics achievement of 4th grade students differently based on sex and ethnicity of students?

H₀: Ability grouping by mathematics achievement of 4th grade students does not affect students differently by sex and ethnicity.

H₁: Ability grouping by mathematics achievement of 4th grade students affects students differently by sex and ethnicity.

Connection to Back to Literature Review

The field of educational research focused on mathematics equity has seen increased scrutiny over the past twenty years. Secada, Fennema, and Adajian (1995) warned that the urgency of researchers to look into issues of equity has led to a rush for answers and solution (p. 149). Lubienski (2002) argued that “although researchers and reformers give attention to equity, such work tends to ignore relevant social and cultural issues” (p, 103). A past review of mathematics research shows limited emphasis on classroom processes and ethnicity, class, and sex in relation to student cognition (Lubienski, 2002, p. 107). Lubienski states that most researchers of mathematics instruction are unfamiliar with theories relating to culture and power. A background in anthropology and sociology is just as important as that of mathematics pedagogy. Mathematics instruction research with a diversity focus is both marginalized and underdeveloped (Cobb & Hodge, 2002). Additional research is needed to explore the systematic and structural aspects of inequity in mathematics instruction (Lester Jr, 2007).

A review of literature exposed a lack of current and past research on not only the antecedents of grouping mathematics students in elementary school but the effects that grouping has on these students.

There is also currently a lack of research into the structural processes that upper elementary school teachers use to recommend students for advanced mathematics courses and the affects of grouping elementary students into homogenous mathematics groups.

Methodology

This study used a causal-comparative design in an attempt to find relationships between independent and dependent variables (see Appendix B). A causal-comparative design “seeks to find relationships between independent and dependent variables after an action or event has already occurred” (Salkind, 2010). The independent variables were student sex, student ethnicity, whether or not students were grouped into mathematics ability groups, and the level of mathematics ability group. The dependent variable of this study was the overall change in mathematics achievement from fall to spring. The goal of this quantitative study was to determine whether the independent variable is related to the outcome variable by comparing two or more groups of individuals after the event has occurred (*ex post facto*).

A causal-comparative design is frequently used within educational research where independent variables are already fixed in place (Salkind, 2010). Classrooms are not re-organized based on the experimental design constraints of a research study. Another key aspect to causal-comparative designs is subjects are already in groups in contrast to a true experimental research design where subjects are randomly selected (Salkind, 2010).

According to Salkind (2010), inferential statistical methods, such as chi-square test, paired-samples and independent t tests, and analysis of variance (ANOVA), are appropriate within a casual-comparative research study when the “researcher hopes to demonstrate that a relationship exist between the independent and dependent variables.”

Crosstabs and chi-square test are used to compare two nominal variables, a nominal variable with an ordinal variable, or two ordinal variables (Muijs, 2010). The distribution of females and males across the student categories of general education, special education, and limited English proficiency were explored using crosstabs. The distribution of ethnicity across student categories was also evaluated. The statistical significance of these distributions was analyzed using chi-square test.

T-tests are appropriate when comparing the differences between the means of a continuous variable between two groups (Muijs, 2010). Two different continuous dependent variables were analyzed. The first is the change in the student's scaled score (change_SS) that was measured from their fall STAR mathematics assessment and that of their spring score. The second dependent variable is change in a student's grade equivalency (change_GE) that was measured during the same time frame as their scaled score. Independent sample t-test were conducted to check the statistical significance and effect size between the dependent variable (mathematics achievement) and the independent variables of sex and whether or not students were grouped in mathematics.

Analysis of variance (ANOVA) allows comparison of a continuous dependent variable and several groups (Muijs, 2010). An F-test can be used post hoc to determine not only the statistical significance between groups, but the effect size. Main effects and interactions were analyzed on the scale score change (change_SS) dependent variable and the independent variables of sex and ethnicity.

Ideally, a multiple linear regression would be used to predict the change in scale score (change_SS) based on a student's sex, ethnicity, and math grouping. This was unable to be accomplished given the lack of representation as students were put in

categories based on their sex, ethnicity, and grouping level. There was only sufficient representation to analysis sex and grouping level versus the change_SS using a linear regression model. This will help in the construction of a linear equation that will predict values of the dependent variable based on the independent variables (Gray & Kinnear, 2012).

Interpretation of Findings

Grouping

The primary focus of this study was to look into the effects that homogeneously grouping mathematics students based on their ability had on their mathematics achievement over the course of 4th grade. The change in mathematics achievement was investigated across mathematics ability groups, ethnicity, and sex.

The first research question, “Does ability grouping affect mathematics achievement of 4th grade students?”, was analyzed with both an independent samples t-test and a Mann-Whitney U test. Both of these analyses showed that even though the grouped mathematics students had a higher growth, this change in scaled score (change_SS) and change in grade equivalency (change_GE) was not statistically significant. The change in grade equivalency (change_GE) did not pass the assumption of normality and the assumption of homogeneity of variance. In this case, a nonparametric test that requires fewer assumptions of equality is appropriate (Gray & Kinnear, 2012).

In regards to the dichotomous variable of “whether or not students are grouped”, the null hypothesis should not be rejected. At least for this study’s sample, mathematics achievement was not dependent on whether or not students were placed in either homogenous or heterogeneous ability groups. Hattie’s meta-analyses (2008) found

similarly that tracking not only has minimal effects on learning outcomes but has profound negative equity effects.

An analysis of variance (ANOVA) was conducted to determine whether there was a significant difference in growth between each of the four groups (below grade level = low, at grade level = medium, above grade level=high, heterogeneous grouped = mixed.) The result of this ANOVA was counterintuitive. A main effect of grouping on change in scaled score (change_SS) was statistically significant. Post-hoc analyses using Tukey's HSD indicated that the change_SS had a statistically lower mean ($p = .005$) for the medium grouped students ($N=79$, $M=63.73$, $SD=60.517$) than the low grouped students ($N=50$, $M=101.12$, $SD=56.033$) (Table 6: Change_SS vs Levels of Grouping). The model weakly accounts for the variance in the change in a student's scaled score based solely on how they are grouped ($R^2 = 0.024$).

The implication for this finding is that the lower grouped students outperformed the medium grouped students within this study's sample. Several threats to the internal validity, as defined by Krathwohl (1985), can help to explain why the lower grouped students performed better than the medium grouped students. The first is a *history threat*. Students in the lower group may be receiving more mathematics instruction in terms of time. Students in the lower group could be receiving additional mathematics interventions that the medium grouped students are not receiving. More resources could be allocated for a population of students that is viewed, at least at the beginning of the school year, as being at risk in terms of mathematical understanding. Another threat could be a combined *testing threat* and *expectancy effect*. It is plausible, but not accounted for within the confines of this study, that the lower grouped students were tested more routinely using

the STAR assessment and thus had more experience with the assessment than the medium grouped students. Additionally, the teachers and paraprofessionals that work with the lower grouped students could be either implicitly or explicitly providing higher motivating signals to their students.

Ethnicity

The second research question, “Does ability grouping affect differently the mathematics achievement of 4th grade students based on the sex and ethnicity of students?” was analyzed with an ANOVA. Ethnicity could only be analyzed across sex. Lack of ethnic representation across all grouping levels prevented the comparison of change_SS based on a student’s ethnicity and grouping level.

Hispanic students performed statistically worse than both Asian or Pacific Islander students and White students within this study’s sample. The performance of Hispanic students within this study is similar to data from the National Assessment of Educational Progress (NAEP) that showed from 1990 to 2009, the achievement gap between fourth-grade Hispanic and White public-school students remained at around 21-points. More alarming is that this gap increases to 26-points for eighth-grade Hispanic and Whites over the same time frame (Hemphill & Vanneman, 2011).

Not only did Hispanic students perform lower than other students, they were not represented within the above grade level group. Of the seventy-five Hispanic students within this study, zero were members of the above grade level group, which had forty-three students. Sixty-four Hispanic students were in the mixed group, 4 in the low group, and 7 in the medium group. This lack of representation within the above grade level group is characteristic of what past literature has stated. Grouping and tracking students

inevitably separates students by race, ethnicity, native language, and class (Loveless & Diperna, 2000). In the past, Latino students who scored within the 60th percentile were 50 percent less likely to be placed into a college preparatory class when compared to White and Asian students (Oakes, 1992). Not only are non-Asian minority and poor students tracked into lower mathematics courses, but those courses tend to have less qualified teachers (Secada, 1992). This trend is still true today, with Black, Hispanic and poor children still frequently tracked into remedial classes, while middle-class White children are tracked into honors courses (Loveless & Diperna, 2000).

Sex

Female students performed statistically worse than male students within this study's sample population. Female students, as a group, scored almost 15 points less than their male counterparts (see Table 19: Sex vs Change_SS). Though female students were equally represented in all four mathematics groups, they performed lower than the males in each of the four groups (see Table 20: Sex and Grouping vs Change_SS).

These findings are consistent with the results from National Assessment of Educational Progress (NAEP) in the years 1990 to 2003 (McGraw, Lubienski, & Strutchens, 2006). Throughout that reporting period, 4th grade girls lagged behind boys in all areas of mathematical reasoning.

A 4x2 ANOVA with sex (female and male) and grouping level (mixed, low, medium, high) as between-subjects' factors found there was no significance between the interaction of sex and grouping.

One possible rival explanation as to why female students did not perform as well as male students could be attributed to an instrumentation threat to internal validity

(Krauthwohl, 1985). Though the STAR assessment was normed using a representative sample of both female and male students (Learning, 2015), internal validity scores were only published based on grade level (Learning, 2014). It is possible that females are reacting to a sex bias inherent within the assessment.

Another explanation to the sex differential within this study's sample is raised by Boaler (2015). She states elementary school girls identify with their female teachers and quickly pick up on their teachers' negative messages about math. This assertion is based on Beilock, Gunderson, Ramirez, and Levine's (2010) work on how the math anxiety of female teachers affects their female students.

Ethnicity, Sex, and Grouping

An analysis of variance (ANOVA) and a multiple linear regression found no significant interactions between sex and ethnicity, and sex and grouping. Sex and ethnicity were both predictive of mathematics achievement without regard to each other or with grouping taken into account. This collides with Gutiérrez's assertion that mathematics equity is the "inability" to predict students' mathematics achievement and participation based solely on race, class, ethnicity, sex, beliefs, and language proficiency (Gutiérrez & Dixon-Román, 2010). At least for the participants within this study, both sex and ethnicity were predicative, though weakly, of students' mathematical achievement.

Practical Implication for Educators

Educators and school administrators at all levels want the best for their students. It is within this mindset that elementary school instructional teams go about trying to structure their mathematics classrooms to maximize the achievement of each student.

Unfortunately, this study helps to corroborate what past research has found, that grouping has minimal effects on learning outcomes (Hattie, 2008). Elementary school staff and district administrators feel the weight of trying to attenuate the achievement gap at the cost of the opportunity gap. At times strong advocacy for homogenous grouping of mathematics students within elementary schools comes from “high-track” teachers and parents who have benefited from tracking (Mathis, 2013).

This study highlights the need for the continuous review of mathematics performance not only at the classroom level, but at the grade-level based on sex and ethnicity. Sub-group (sex, ethnicity, SES, etc...) level growth performance is not discernable at the classroom level given the low sample size. For example, teachers and school administrators may not be able to detect that females are falling further behind their male counterparts.

This study has also put a spotlight on the need to address the mathematical reasoning of females and Hispanic students. McGraw, Lubienski, and Strutchens (2006) state that sex differences within mathematics achievement can be related to students’ attitudes about mathematics, their self-concepts, their self-confidence, and teacher beliefs, expectations, and interactions with male and female students.

The differences in achievement between Hispanic students and other groups will only become more significant as time goes on. Recent predications state that the Hispanic population will increase from 17% of the population in 2014 to 29% of the population by 2060 (Colby & Ortman, 2015). Abreu and Cline’s (2003) work on social valorization of mathematics practices may help to bring school level discourse to the issue of mathematics equity across all ethnic groups.

Recommendations for Further Study

The generalizability of this study is constrained due to its limited and non-randomized sample. In light of that, several opportunities exist for further study at the district level and ultimately on a broader scale.

NEAP data suggest that the mathematics achievement differences across sex and ethnicity increase as students increase in grade level (McGraw, Lubienski, & Strutchens, 2006). Are the gaps in mathematics achievement of both females and Hispanic students limited to just 4th grade or do they persist throughout these populations across all grade levels within this school district?

Another area for further study would be to determine whether female and Hispanic students are underrepresented in advanced mathematics opportunities within the school district represented in this study. Advanced mathematics opportunities could be defined as Advance Placement (AP) courses and enrollment in the district's self-contained mathematics and science secondary school.

This study did not attempt to study the antecedents of grouping 4th grade students. Areas of opportunity would be to study 1) what factors are used when students are grouped into mathematics ability groups within elementary schools, 2) what determines which teacher will teach which level of mathematics within elementary schools, and 3) at what grade level does sorting begin in mathematics?

Conclusion

This study highlights the need for continuous research and reflection in all areas of mathematics pedagogy and equity. Teachers and administrators come to school each day with the desire to better their students' lives. It is with a conviction of having a

growth mindset that the school district participated with this study. The results of this study will be one point on a continuum of self-improvement and deliberation.

REFERENCES

- Abreu, G. (1995). Understanding how children experience the relationship between home and school mathematics. *Mind, Culture, and Activity*, 2(2), 119-142.
- Abreu, G. (2007). Social valorisation of mathematics practices: the implications for learners in multicultural schools. In Nasir, N. & Cobb, P. (Eds.) *Improving Access to Mathematics: Diversity and Equity in the Classroom*. (pp. 118-131) Teachers College Press.
- Abreu, G. D., & Cline, T. (2003). Schooled mathematics and cultural knowledge. *Pedagogy, Culture and Society*, 11(1), 11-30.
- Adelman, C. (1999). *Answers in the Tool Box. Academic Intensity, Attendance Patterns, and Bachelor's Degree Attainment*.
- Adelman, C. (2006). *The Toolbox Revisited: Paths to Degree Completion From High School Through College*. US Department of Education.
- Akiba, M., LeTendre, G. K., & Scribner, J. P. (2007). Teacher quality, opportunity gap, and national achievement in 46 countries. *Educational Researcher*, 36(7), 369-387.
- Althusser, L. (2006). Ideology and ideological state apparatuses (notes towards an investigation). *The anthropology of the state: A reader*, 9, 86.
- Apple, M. W. (2004). *Ideology and curriculum*. Routledge.
- Apple, M. W., & Franklin, B. (2004). Curricular history and social control. *Ideology and curriculum*, 59-76.
- Bartell, T. G. (2007). Culture, race, power, and mathematics instruction. *Second handbook of research on mathematics teaching and learning*, 405-434.

- Battey, D. (2013). Access to Mathematics: "A Possessive Investment in Whiteness" 1. *Curriculum Inquiry*, 43(3), 332-359.
- Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences*, 107(5), 1860-1863.
- Belenky, M. F. (1986). *Women's ways of knowing: The development of self, voice, and mind*. Basic books.
- Benbow, C. P., & Stanley, J. C. (1996). Inequity in equity: How "equity" can lead to inequity for high-potential students. *Psychology, Public Policy, and Law*, 2(2), 249.
- Boaler, J. (2015). *Mathematical mindsets: Unleashing students' potential through creative math, inspiring messages and innovative teaching*: John Wiley & Sons.
- Bohrnstedt, G., Kitmitto, S., Ogut, B., Sherman, D., & Chan, D. (2015). *School Composition and the Black-White Achievement Gap*. NCES 2015-018. National Center for Education Statistics.
- Bond, G., & Chernoff, E. J. (2015). Mathematics and Social Justice: A Symbiotic Pedagogy. *Journal of Urban Mathematics Education*, 8(1), 24-30.
- Bose, E., & Remillard, J. (2010). Looking for Equity in Policy Recommendations for Instructional Quality. In *Mapping Equity and Quality in Mathematics instruction* (pp. 177-190). Springer Netherlands.
- Bourdieu, P. (1977). *Outline of a Theory of Practice* (Vol. 16). Cambridge university press.
- Bruner, J. S. (1986). *Actual minds, possible worlds*. Harvard University Press.
- Burris, C. C., Welner, K. G., & Bezoza, J. W. (2009). *Universal access to a quality education: Research and recommendations for the elimination of curricular stratification*: Education and the Public Interest Center, School of Education, University of Colorado.

- Carnevale, A. P., & Desrochers, D. M. (2003). The democratization of mathematics. Quantitative literacy: Why numeracy matters for schools and colleges, 21-31.
- Cobb, P., & Hodge, L. L. (2002). A relational perspective on issues of cultural diversity and equity as they play out in the mathematics classroom. *Mathematical thinking and learning*, 4(2-3), 249-284.
- Cobb, P., & Hodge, L. L. (2010). Culture, identity, and equity in the mathematics classroom. In *A journey in mathematics instruction research* (pp. 179-195). Springer Netherlands.
- Colby, S. L., & Ortman, J. M. (2015). Projections of the Size and Composition of the US Population: 2014 to 2060. *US Census Bureau, Ed*, 25-1143.
- Cornell, S., & Hartmann, D. (2006). Ethnicity and race: Making identities in a changing world. Sage Publications. curriculum, 59-76.
- D'Ambrosio, B., Frankenstein, M., Gutierrez, R., Kastberg, S., Martin, D. B., Moschkovich, J., . . . Barnes, D. (2013). Introduction to the JRME Equity Special Issue JRME Equity Special Issue Editorial Panel. In: NATL COUNC TEACH MATH 1906 ASSOCIATION DRIVE, RESTON, VA 22091 USA.
- D'Amato, J. (1993). Resistance and compliance in minority classrooms. *Minority education: Anthropological perspectives*, 181-207.
- Darling-Hammond, L. (2010). *The flat world and education: how America's commitment to equity will determine our future*. New York: Teachers College Press.
- Dornyei, Z. (2007). *Research methods in applied linguistics*.
- El Nasser, H. (2004). Census projects growing diversity. By 2050: population burst, societal shifts. *USA Today*, 3(18), 04.
- Ernest, P. (2002). Empowerment in mathematics instruction. *Philosophy of mathematics instruction journal*, 15(1), 1-16.
- Evan, A., Gray, T., & Olchefske, J. (2006). *The Gateway to Student Success in Mathematics and Science: A Call for Middle School Reform; The Research and Its Implications*. Washington, DC: American Institutes for Research

- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, 41(4), 1149-1160.
- Flores, A. (2007). Examining disparities in mathematics instruction: Achievement gap or opportunity gap?. *The High School Journal*, 91(1), 29-42.
- Freire, P. (2000). *Pedagogy of the oppressed*. Bloomsbury Publishing.
- Fry, R. (2007). *The Changing Racial and Ethnic Composition of US Public Schools*. Pew Hispanic Center.
- Goldhaber, D., Lavery, L., & Theobald, R. (2015). Uneven playing field? Assessing the teacher quality gap between advantaged and disadvantaged students. *Educational Researcher*, 0013189X15592622.
- González, N., Andrade, R., Civil, M., & Moll, L. (2001). Bridging funds of distributed knowledge: Creating zones of practices in mathematics. *Journal of Education for students placed at risk*, 6(1-2), 115-132.
- Gordon, R. J., Kane, T. J., & Staiger, D. (2006). *Identifying effective teachers using performance on the job*. Washington, DC: Brookings Institution.
- Gray, C. D., & Kinnear, P. R. (2012). *IBM SPSS statistics 19 made simple*: Psychology Press.
- Greenberg, J., & Walsh, K. (2008). *No Common Denominator: The Preparation of Elementary Teachers in Mathematics by America's Education Schools*. National Council on Teacher Quality.
- Gutiérrez, R. (2007). defining equity: The importance of a critical perspective. *Improving access to mathematics: Diversity and equity in the classroom*, 37-50.
- Gutiérrez, R., & Dixon-Román, E. (2010). Beyond Gap Gazing: How Can Thinking About Education Comprehensively Help Us (Re) envision Mathematics instruction?. In *Mapping equity and quality in mathematics instruction* (pp. 21-34). Springer Netherlands.

- Gutstein, E. (2007). So one question leads to another”: Using mathematics to develop a pedagogy of questioning. *Improving access to mathematics: Diversity and equity in the classroom*, 51-68.
- Haberman, M. (1991). The pedagogy of poverty versus good teaching. *Phi Delta Kappan*, 73(4), 290-294.
- Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*: Routledge.
- Hemphill, F. C., & Vanneman, A. (2011). *Achievement Gaps: How Hispanic and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress. Statistical Analysis Report. NCES 2011-459. National Center for Education Statistics.*
- Hoogeveen, L., Van Hell, J. G., & Verhoeven, L. (2005). Teacher attitudes toward academic acceleration and accelerated students in the Netherlands. *Journal for the Education of the Gifted*, 29(1), 30-59.
- Howard, T. C. (2010). *Why Race and Culture Matter in Schools: Closing the Achievement Gap in America's Classrooms. Multicultural Education Series. Teachers College Press.*
- IBM Corp. Released 2016. *IBM SPSS Statistics for Macintosh, Version 24.0. Armonk, NY: IBM Corp.*
- Irvine, J. J., & York, E. (1993). Differences in teacher attributes of school failure for African American, Hispanic, and Vietnamese students.
- Jackson, T. F. (2013). *From civil rights to human rights: Martin Luther King, Jr., and the struggle for economic justice. University of Pennsylvania Press.*
- Jordan, H., Mendro, R., & Weerasinghe, D. (1997). *Teacher effects on longitudinal student achievement: A report on research in progress. National Evaluation Institute.*

- Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and social psychology review*, 9(2), 131-155.
- Krathwohl, D. R. (1985). *Social and behavioral science research : a new framework for conceptualizing, implementing, and evaluating research studies* (1st ed.). San Francisco: Jossey-Bass.
- Ladson-Billings, G. (1995). Making mathematics meaningful in multicultural contexts. *New directions for equity in mathematics instruction*, 126-145.
- Ladson-Billings, G. (1997). It doesn't add up: African American students' mathematics achievement. *Journal for Research in Mathematics instruction*, 28(6), 697-708.
- Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in US schools. *Educational researcher*, 35(7), 3-12.
- Lave, J. (1991). Situating learning in communities of practice. *Perspectives on socially shared cognition*, 2, 63-82.
- Learning, R. (2014). *The research foundation for STAR Assessments: The science of STAR*. Wisconsin Rapids, WI.
- Learning, R. (2015). *STAR math: Technical manual*. Wisconsin Rapids, WI: Renaissance Learning.
- Lester Jr, F. K. (Ed.). (2007). *Culture, race, power, and mathematics instruction*. Second handbook of research on mathematics teaching and learning, 1, 405-433.
- Lewis, A. (2003). *Race in the schoolyard: Negotiating the color line in classrooms and communities*. Rutgers University Press.
- Loveless, T., & Diperna, P. (2000). *How well are American students learning? : Brown Center on Education Policy*, The Brookings Institution.
- Lubienski, S. T. (2002). Research, reform, and equity in US mathematics instruction. *Mathematics thinking and learning*, 4(2-3), 103-125.
- Lubienski, S. T. (2008). On " gap gazing" in mathematics instruction: The need for gaps analyses. *Journal for Research in Mathematics instruction*, 350-356.

- Martin, D. (2007). Mathematics learning and participation in African American context: The co-construction of identity in two intersecting realms of experience. *Diversity, equity, and access to mathematics ideas*, 146-158.
- Martínez, J. F., Stecher, B., & Borko, H. (2009). Classroom assessment practices, teacher judgments, and student achievement in mathematics: Evidence from the ECLS. *Educational Assessment*, 14(2), 78-102.
- Mathematics, N. C. T. M. (2000). Principles and standards for school mathematics (Vol. 1): National Council of Teachers of.
- Mathis, W. (2013). Moving beyond tracking. Boulder, CO: National Education Policy Center. Retrieved July, 3, 2013.
- Mayer, D. P., Mullens, J. E., & Moore, M. T. (2000). Monitoring school quality: An indicators report (NCES 2001-030). US Department of Education. National Center for Education Statistics. Washington, DC: US Government Printing Office.
- McGraw, R., Lubienski, S. T., & Strutchens, M. E. (2006). A Closer Look at Gender in NAEP Mathematics Achievement and Affect Data: Intersections with Achievement, Race/Ethnicity, and Socioeconomic Status. *Journal for Research in Mathematics Education*, 37(2), 129-150. doi:10.2307/30034845
- McMillan, J. H. (1996). Educational research: Fundamentals for the consumer: ERIC.
- Moses, R. P., & Cobb, C. E. (2001). Radical equations: math literacy and civil rights. Boston: Beacon Press.
- Muijs, D. (2010). Doing quantitative research in education with SPSS: Sage.
- Musu-Gillette, L., Robinson, J., McFarland, J., KewalRamani, A., Zhang, A., & Wilkinson-Flicker, S. (2016). Status and Trends in the Education of Racial and Ethnic Groups 2016. NCES 2016-007. *National Center for Education Statistics*.
- Nasir, N. I. S. (2002). Identity, goals, and learning: Mathematics in cultural practice. *Mathematics thinking and learning*, 4(2-3), 213-247.

- National Center for Education Statistics, United States Department of Education. (2002). Defining race and ethnicity data. National Center for Education Statistics. Retrieved from http://nces.ed.gov/statprog/2002/std1_5.asp
- National Council of Teachers of Mathematics (Ed.). (2000). Principles and standards for school mathematics (Vol. 1). National Council of Teachers of.
- National Mathematics Advisory Panel. (2008). Foundations for success: The final report of the National Mathematics Advisory Panel. US Department of Education.
- NCES. (2016). Limited English Proficient (LEP) Students in Limited English Proficient Program - School Years 2011-12 and 2012-13. Washington, D.C.
- Nulty, D. D. (2008). The adequacy of response rates to online and paper surveys: what can be done? *Assessment & evaluation in higher education*, 33(3), 301-314.
- Nunley, K. F. (2004). Layered curriculum: *The practical solution for teachers with more than one student in their classroom*. Kathie F. Nunley.
- Oakes, J. (1985). Keeping track. Yale University Press.
- Oakes, J. (1987). Tracking in secondary schools: A contextual perspective. *Educational Psychologist*, 22(2), 129-153.
- Oakes, J. (1990). Multiplying Inequalities: The Effects of Race, Social Class, and Tracking on Opportunities to Learn Mathematics and Science.
- Oakes, J. (1992). Can tracking research inform practice? Technical, normative, and political considerations. *Educational Researcher*, 21(4), 12-21.
- Packer, L. (2002). Special education glossary and acronyms guide. Retrieved October, 23, 2011.
- Pais, A., & Valero, P. (2010). Beyond disavowing the politics of equity and quality in mathematics instruction. In Mapping equity and quality in mathematics instruction (pp. 35-48). Springer Netherlands.
- PISA, O. E. C. D. (2012). Results: What Students Know and Can Do: Student Performance in Mathematics. *Reading and Science*, 1.

- Planas, N., & Civil, M. (2009). Working with mathematics teachers and immigrant students: An empowerment perspective. *Journal of Mathematics Teacher Education*, 12(6), 391-409.
- Qualtrics. (2017). Qualtrics. Retrieved March 2017 <http://www.qualtrics.com/>
- Reed, R. J., & Oppong, N. (2005). Looking critically at teachers' attention to equity in their classrooms. *The Mathematics Educator*, 1, 2-15.
- Remillard, J. (2000). Prerequisites for Learning to Teach Mathematics for All Students. *Changing the faces of mathematics instruction: perspectives on multiculturalism and gender equity*, Reston, VA, National Council of Teachers of Mathematics, 37-46.
- Rich, A. (1979). Invisibility in academe. *Blood, bread, and poetry: Selected prose 1979-1985*, 198-201.
- Rose, S., Spinks, N., & Canhoto, A. (2015). Tests for the assumption that a variable is normally distributed. *Management research: Applying the principles*.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectations and student intellectual development*.
- Rousseau, C., & Tate, W. F. (2003). No time like the present: Reflecting on equity in school mathematics. *Theory into Practice*, 42(3), 210-216.
- Roy, F. (2000). Technology and equity: a consistent vision for school mathematics?'. *Changing the faces of mathematics instruction: perspectives on multiculturalism and gender equity*, Reston, VA, National Council of Teachers of Mathematics, 37-46.
- Salkind, N. J. (2010). *Encyclopedia of research design* (Vol. 1): Sage.
- Schoenfeld, A. H. (2002). Making mathematics work for all children: Issues of standards, testing, and equity. *Educational researcher*, 31(1), 13-25.
- Secada, W. G. (1989). Agenda setting, enlightened self-interest, and equity in mathematics instruction. *Peabody Journal of Education*, 66(2), 22-56.

- Secada, W. G. (1992). Race, ethnicity, social class, language, and achievement in mathematics.
- Secada, W. G., Fennema, E., & Byrd Adajian, L. (1995). *New directions for equity in mathematics education*. Cambridge ; New York, NY, USA: Cambridge University Press.
- Silver, E. A., Smith, M. S., & Nelson, B. S. (1995). The QUASAR project: Equity concerns meet mathematics instruction reform in the middle school. *New directions for equity in mathematics instruction*, 9-56.
- Stanic, G. M. (1989). Social inequality, cultural discontinuity, and equity in school mathematics. *Peabody Journal of Education*, 66(2), 57-71.
- Stevens, T., Olivarez, A., Lan, W. Y., & Tallent-Runnels, M. K. (2004). Role of mathematics self-efficacy and motivation in mathematics performance across ethnicity. *The Journal of Educational Research*, 97(4), 208-222.
- Stinson, D. W. (2004). Mathematics as “gate-keeper”(?): Three theoretical perspectives that aim toward empowering all children with a key to the gate.
- Strutchens, M. E. (2000). Confronting beliefs and stereotypes that impede the mathematics empowerment of African American students. *Changing the faces of mathematics: Perspectives on African Americans*, 7-14.
- Tate, W. F. (2008). The Political Economy of Teacher Quality in School Mathematics African American Males, Opportunity Structures, Politics, and Method. *American Behavioral Scientist*, 51(7), 953-971.
- Thompson, A. (1998). Not the color purple: Black feminist lessons for educational caring. *Harvard Educational Review*, 68(4), 522-555.
- Useem, E. L. (1992). Middle schools and math groups: Parents' involvement in children's placement. *Sociology of education*, 263-279.
- Valero, P. (2009). Mathematics instruction as a network of social practices. In Congress of the European Society for Research in Mathematics instruction.

- Velez-Ibanez, C. G., & Greenberg, J. B. (1992). Formation and Transformation of Funds of Knowledge among U.S.-Mexican Households. *Anthropology and Education Quarterly*, 23, 4, 313-35.
- Wells, C. S., & Wollack, J. A. (2003). An instructor's guide to understanding test reliability. *Testing & Evaluation Services. University of Wisconsin.*
- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity.* Cambridge university press.
- Wiener, R. (2006). Opportunity gaps: The injustice underneath achievement gaps in our public schools. *NCL Rev.*, 85, 1315.
- Žižek, S. (2009). *The parallax view.* M.I.T. Press.

APPENDIX A

4th Grade Mathematics Grouping Online Teacher Questionnaire

Q1.1 4th Grade Mathematics Grouping

Q1.2 Thank you for taking the time to take this survey. It should take approximately 15 minutes. This survey will help to explore how 4th grade students are grouped for mathematics instruction. This study involves no foreseeable serious risks and offers the potential to further understand the effects of grouping within 4th grade mathematics classrooms. Your participation only requires the completion of an on-line survey. This survey will take approximately 15 minutes to complete. The survey is focused on how your students are grouped and your views on grouping 4th grade mathematics students. Student demographics and achievement data will be linked to your survey response to explore the role grouping has in mathematics achievement. Your identity will only be seen by the principal investigator and the co-investigator. Pseudonyms will be assigned to your survey responses. School district, school, and teacher identifications will not be published. By consenting to take this survey you indicate that you understand the purpose of this survey and voluntarily agree to participate. By giving consent to participate in this study, you confirm that the researcher has explained the elements of informed consent. You acknowledge participation is voluntary and do not have to participate. The purpose of the research as well as the risks and benefits have been explained. The procedures as well as the time commitment have been outlined and that you understand the issues of confidentiality. If you have any questions or concerns, feel free to

contact the principal investigator. Questions may also be sent to my faculty supervisor. This study falls under IRB number 108-SB17-032.

Thank you for your help!

Q1.3 Do you give consent to participate in this survey?

- Yes
- No

Q2.1 How many 4th grade homeroom classes are there in your school?

- 1
- 2
- 3
- 4
- 5
- 6
- more than 6

Q2.2 Does your school use the STAR mathematics assessment for 3rd and 4th grade students?

- Yes
- No

Q2.3 Does your school's 3rd and 4th grade mathematics students take the STAR mathematics assessment in both the Fall and Spring?

- Yes
- No

Q2.4 How effective do you believe the STAR mathematics assessment is in measuring math proficiency?

- Not effective at all
- Slightly effective
- Moderately effective
- Very effective
- N/A (We do not use STAR for mathematics assessments.)

Q2.5 Do you use any other math assessment tools with your 3rd and 4th grade mathematics students? Please explain if "Yes".

- Yes _____
- No

Q2.6 What is your current teaching assignment? (Check all that apply)

- New to Profession
- New to District
- General Education
- Gifted and Talented
- Montessori
- Dual Language
- 3rd/4th Grade Combination
- 4th/5th Grade Combination

Q2.7 How many years of work experience do you have? Count this school year as a full year. (Please round up to whole years)

_____ Year(s) teaching 4th grade mathematics

_____ Year(s) working as a teacher at this school

_____ Year(s) working as a teacher in total

_____ Year(s) working in other education roles (do not include years working as a teacher

_____ Year(s) working in other jobs not as a teacher and not in an education role

Q2.8 What is the highest level of formal education you have completed?

- Associate's degree
- Bachelor's degree
- Master's degree
- Doctoral degree or equivalent (Ph.D., Ed.D., J.D., M.D.)

Q2.9 How did you receive your teaching license?

- Formal teacher preparation program
- American Board for Certification of Teacher Excellence (ABCTE)
- Teach for America
- Other alternate path to licensure

Q3.1 How many separate math classes do you teach that includes 4th graders?

- 1
- 2
- 3
- 4

Q3.2 How many days a week do you teach your 4th grade mathematics class?

- 1 day a week
- 2 days a week
- 3 days a week
- 4 days a week
- 5 days a week

Q3.3 How many minutes a day do you teach your 4th grade mathematics class?

- 30 to 40 minutes per day
- 41 to 50 minutes per day
- 51 to 60 minutes per day
- 61 to 70 minutes per day
- Greater than 70 minutes per day

Q3.4 In regards to the mathematics class you teach, to what extent do you feel prepared for the items below? Please mark one choice in each row.

	Not at all	Somewhat	Well	Very Well
Content of the mathematics concepts I teach.....	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pedagogy of the mathematics concepts I teach.....	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Classroom practice of the mathematics concepts I teach.....	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q3.5 In regards to the mathematics class you teach, to what extent do you feel you have the resources you need? Please mark one choice in each row.

	Not at all	Somewhat	Well	Very Well
Curricular resources needed for the mathematics class I teach.....	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Professional development opportunities needed for the mathematics class I teach.....	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Collaboration opportunities needed for the mathematics class I teach.....	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q4.1 Is the mathematics class you teach to your 4th grade students an ability-grouped mathematics class or a mixed-ability mathematics class?

- I teach a low-ability grouped (i.e. below grade level) mathematics class. The majority, over 50%, of my mathematics class for the 2016-2017 school year is below grade level.
- I teach a medium-ability grouped (i.e. at grade level) mathematics class. The majority, over 50%, of my mathematics class for the 2016-2017 school year is at grade level.
- I teach a high-ability grouped (i.e. above grade level) mathematics class. The majority, over 50%, of my mathematics class for the 2016-2017 school year is above grade level.
- I teach a mixed-ability mathematics class. The students in my mathematics class for the 2016-2017 school year have not been grouped by ability level.

Q5.1 I feel that the students in my low-ability grouped mathematics class are in the following grade performance range: (Select all grade levels within range of this class.)

- K
- 1ST
- 2ND
- 3RD
- 4TH
- 5TH
- 6TH
- 7TH
- 8TH
- 9TH
- 10TH
- 11TH
- 12TH

Q5.2 Based on how your mathematics students are grouped in your low-ability grouped class, are they improving their mathematical thinking?

- Below my expectations
- About what I'd expect
- Above my expectations

Q5.3 In general, do you believe that grouping vs. not grouping low-ability mathematics students improves their mathematical thinking?

- Not at all
- Slightly improves
- Moderately improves
- Improves very much

Q6.1 I feel that the students in my current mathematics class are in the following grade performance range: (Select all grade levels within range of this class.)

- K
- 1ST
- 2ND
- 3RD
- 4TH
- 5TH
- 6TH
- 7TH
- 8TH
- 9TH
- 10TH
- 11TH
- 12TH

Q6.2 Based on how your mathematics students are grouped in your medium-ability grouped class, are they improving their mathematical thinking?

- Below my expectations
- About what I'd expect
- Above my expectations

Q6.3 In general, do you believe that grouping vs. not grouping medium-ability mathematics students improves their mathematical thinking?

- Not at all
- Slightly improves
- Moderately improves
- Improves very much

Q7.1 I feel that the students in my current mathematics class are in the following grade performance range: (Select all grade levels within range of this class.)

- K
- 1ST
- 2ND
- 3RD
- 4TH
- 5TH
- 6TH
- 7TH
- 8TH
- 9TH

10TH

11TH

12TH

Q7.2 Based on how your mathematics students are grouped in your high-ability grouped class, are they improving their mathematical thinking?

Below my expectations

About what I'd expect

Above my expectations

Q7.3 In general, do you believe that grouping vs. not grouping high-ability mathematics students improves their mathematical thinking?

Not at all

Slightly improves

Moderately improves

Improves very much

Q8.1 I feel that the students in my current mathematics class are in the following grade performance range: (Select all grade levels within range of this class.)

K

1ST

2ND

3RD

4TH

5TH

6TH

- 7TH
- 8TH
- 9TH
- 10TH
- 11TH
- 12TH

Q8.2 Based on how your mathematics students, who are not grouped by ability, are they improving their mathematical thinking?

- Below my expectations
- About what I'd expect
- Above my expectations

Q8.3 In general, do you believe that grouping vs. not grouping high-ability mathematics students improves their mathematical thinking?

- Not at all
- Slightly improves
- Moderately improves
- Improves very much

Q8.4 In general, do you believe that grouping vs. not grouping high-ability mathematics students improves their mathematical thinking?

- Not at all
- Slightly improves
- Moderately improves
- Improves very much

APPENDIX B

Independent Variables

1. Ethnicity (NCES, 2002)
 - a. Hispanic or Latino
 - b. American Indian or Alaska Native
 - c. Asian/Native Hawaiian or Other Pacific Islander
 - d. Black or African American
 - e. White
2. Student Educational Category
 - a. General Education (G.E.)
 - b. Special Education (S.P.E.D.)
 - c. Limited English Proficiency (L.E.P.)
 - d. Both Special Education and Limited English Proficiency (S.P.E.D./L.E.P.)
3. Sex
 - a. Female
 - b. Male
4. Teacher Inputs
 - a. Teacher Experience – Continuous ratio variable in years
 - b. Education Level of Teacher – Categorical variable
 - i. Associate's degree
 - ii. Bachelor's degree
 - iii. Master's degree
 - iv. Doctoral degree or equivalent (Ph.D., Ed.D., J.D., M.D.)
 - c. Number of days per week of mathematics instruction

- d. Number of minutes per day of mathematics instruction
- e. Ability Grouping Inputs - Categorical variable coded for the following:
 - i. No Grouping or students heterogeneously grouped
 - ii. Students homogenously grouped into below grade level group
 - iii. Students homogenously grouped into at grade level group
 - iv. Students homogenously grouped into above grade level group

APPENDIX C

Teacher Recruitment Script and Consent

Via Email or In-Person with 4th Grade Teacher:

Hello, my name is Brian Marinelli. I am a doctoral student in the College of Education at Boise State University. I am currently conducting a research study about the effects of grouping 4th grade students on mathematical achievement.

I am contacting you to ask if you would consider participating in my study. I have received school district approval from _____ and _____ (school principal) to contact you directly.

This study involves no foreseeable serious risks and offers the potential to further understand the effects of grouping within 4th grade mathematics classrooms. Your participation only requires the completion of a 25-question on-line survey. This survey will take approximately 15 minutes to complete. The survey is focused on how your mathematics students are grouped and your views on grouping 4th grade mathematics students. Student demographics and achievement data will be linked to your survey response to explore the role grouping has in mathematics achievement. Your identity will only be seen by the principal investigator and the co-investigator. Pseudonyms will be assigned to your survey responses. School district, school, and teacher identifications will not be published.

If you give consent to participate in this study, please reply to this email with the following statement.

I confirm that the researcher has explained the elements of informed consent to the me. I acknowledge participation is voluntary and do not have to participate. The purpose of the research as well as the risks and benefits have been explained. The procedures as well as the time commitment have been outlined. I understand the issues of confidentiality.

The on-line survey will be forwarded to you once your consent is received.

If you have any questions or concerns, feel free to contact the principal investigator. Questions may also be sent to the Boise State University's Institutional Review Board's Human Subjects Coordinator at humansubjects@boisestate.edu.

APPENDIX D

Modified Teacher Recruitment Script and Consent

Via Email or In-Person with 4th Grade Teacher:

Hello, my name is Brian Marinelli. I am a doctoral student in the College of Education at Boise State University. I am currently conducting a research study about the effects of grouping 4th grade students on mathematical achievement.

I am contacting you to ask if you would consider participating in my study. I have received school district approval from _____ and _____ (school principal) to contact you directly.

By completing this survey, you will be entered into a raffle to win one of two \$25 Amazon gift cards.

This study involves no foreseeable serious risks and offers the potential to further understand the effects of grouping within 4th grade mathematics classrooms. Your participation only requires the completion of an on-line survey. This survey will take approximately 15 minutes to complete. The survey is focused on how your mathematics students are grouped and your views on grouping 4th grade mathematics students. Student demographics and achievement data will be linked to your survey response to explore the role grouping has in mathematics achievement. Your identity will only be seen by the principal investigator and the co-investigator. Pseudonyms will be assigned to your

survey responses. School district, school, and teacher identifications will not be published.

If you give consent to participate in this study, please reply to this email with the following statement.

I confirm that the researcher has explained the elements of informed consent to the me. I acknowledge participation is voluntary and do not have to participate. The purpose of the research as well as the risks and benefits have been explained. The procedures as well as the time commitment have been outlined. I understand the issues of confidentiality.

The on-line survey will be forwarded to you once your consent is received.

If you have any questions or concerns, feel free to contact the principal investigator. Questions may also be sent to Boise State University's Institutional Review Board's Human Subjects Coordinator at humansubjects@boisestate.edu. This study falls under IRB number 108-SB17-032.