2021

District Scores and Demographics as Predictors of State Mathematics Assessment Scores

Carlos Alexander Bocel
Walden University

Walden University

College of Social and Behavioral Sciences

This is to certify that the doctoral dissertation by

Carlos Alexander Bocel

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Jesus Tanguma, Committee Chairperson, Psychology Faculty
Dr. Stephen Rice, Committee Member, Psychology Faculty
Dr. Neal McBride, University Reviewer, Psychology Faculty

Chief Academic Officer and Provost
Sue Subocz, Ph.D.

Walden University
2021

Abstract

District Scores and Demographics as Predictors of State Mathematics Assessment Scores

by

Carlos Alexander Bocel

MS, Pace University, 2010

BS, Vaughn College of Aeronautics, 2005

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Educational Psychology

Walden University

February 2021

Abstract

Prior research in mathematical assessments has indicated varying results of predictive variables and further research has been recommended to support students, parents, teachers, and school administration. The purpose of this research was to determine how well a student's performance on a mathematics domain at the state level may be predicted based on student's midyear and end of the year assessment scores, race, gender, and socioeconomic status. The constructivist theoretical foundation was reviewed because of the impact that this theory has on the assessments being researched. Archival records (n = 100) for eighth grade students were received from suburban South Florida charter school and analyzed using multiple linear regression analysis. The results of the multiple linear regression were significant, $F(5, 94) = 32.289$, $p < .001$, and $R^2 = 0.632$. Midyear score ($t = 5.115$, $p < .0001$), and end of year score ($t = 3.92$ $p < .0001$) significantly predicted overall state score. Similarly, midyear score ($t = 2.271=$, $p < .05$), and end of year score ($t = 4.005$, $p < .0001$) significantly predicted the geometry state score, $F(5, 94) = 8.753$, $p < .001$, and $R^2 = 0.318$. Furthermore, the algebra state score was significantly predicted with $F(5, 94) = 19.478$ $p < .0001$, and $R^2 = 5.09$, midyear score ($t = 4.997$, $p < .0001$), and end of year score ($t = 4.493$, $p < .0001$). Finally, midyear score ($t = 3.156$, $p < .05$), and end of year score ($t = 2.449$, $p < .05$) significantly predicted the number sense state score, $F(5, 94) = 6.384$, $p < .0001$, and $R^2 = .254$. Race, gender, and socioeconomic status did not provide predictive value for any of the regression models. These results may have the potential of providing positive social change by adding confidence and support to all stakeholders.

District Scores and Demographics as Predictors of State Mathematics Assessment Scores

by

Carlos Alexander Bocel

MS, Pace University, 2010

BS, Vaughn College of Aeronautics, 2005

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Educational Psychology

Walden University

February 2021

Dedication

I dedicate my work to my daughters Areli and Abigail. Everything I have done is for our daughters. I want them to know that anything is possible with hard work and determination. I hope that through me may be able see the light of God and be light themselves.

Acknowledgments

Education is truly a group effort. I would not have been able to get this completed without several people in my life. I want to first thank my committee chair Dr. Jesus Tanguma. Towards the end we were both working full time on this research.

To my parents, for always supporting me in everything I ever wanted to do. I want to thank everyone gave me words of encouragement when I really needed it. If you are reading this thank you for your support.

I owe so much to my wife Marcela Bocel. She has always been my number one supporter. This path has been years of challenge and struggle, but she has always been there making it easier to get through every obstacle.

Most importantly I want to thank God for giving me life and the opportunity to complete this and many accomplishments. In gratitude, I will spend the rest of my time here helping others fulfil their purpose. I am prepared.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction to the Study

Introduction

The Common Core State Standards (CCSS) were initiated in 2010 with the goal to standardize skills and concepts in specific subjects across the United States (CCSS Initiative, n.d.). Standards were written for K-12 English and mathematics courses and supplement 6-12 history, social studies, science, and technical subjects (CCSS Initiative, n.d.). The CCSS initiatives were intended to create consistent learning goals for students (CCSS Initiative, n.d.). There are four levels of educational administration: the federal, the state, the district, and the individual school level. The order of policy application follows this hierarchal order (CCSS Initiative, n.d.). An individual state decides on how best to proceed with the federal policy. A district then decides how they are to meet the state directives. Lastly, a school's administration is tasked to execute guidelines given by the district, as well as to gather and compare assessment data and how demographics affect student assessment scores to make future decisions (CCSS Initiative, n.d.). It is also important for teachers, parents, and students to know whether assessments are producing the desired results.

Background

Research has been conducted on the mathematics curriculum to indicate whether the students were being prepared to compete globally (Grady et al., 2012). The reasons given for the research were the low results in 2003 and 2007 of the International Mathematics and Science study, as well as the low scores in 2003, 2006, and 2009 of the Program for the International Student Assessment (Grady et al., 2012). In both cases,

students in the United States were being compared with students from other nations. Internal comparison of the National Assessment of Educational Progress indicated that there had been no growth for fourth and eighth grade students in the year 2009 (Grady et al., 2012). In a different study, the U.S. Department of Education received the recommendation that the pedagogy delivery must vary (Grady et al., 2012). This led to the focus on the constructivist approach in curriculum and the creation of current educational standards.

Problem Statement

Henderson et al. (2007) implemented a quasi-experimental design to determine if students who were given quarterly assessments in middle-school mathematics showed greater gains than students who were not given the quarterly assessments. Their study indicated that there was no significant difference between the two samples. Henderson et al. urged researchers to continue to track achievement data to provide current information to policy makers regarding the implications of quarterly assessments. Senator Stephen of Utah underscored the need for good, honest data so Congress could make good policy decisions (Exstrom & Thatcher, 2014).

District and state assessments are based on the current national standards. The assessment questions are constructed to measure the level of student mastery of a standard (CCSS Initiative, n.d.). The article by Main (2012) is a prime example of a detailed account of the emergence of the standards. The author reviewed the timing of introduction of new material, the lack of established curriculum, and the need for professional development to prepare teachers. According to Main, more data collection

and research is recommended to provide a better picture of the standards is also presented. Main advised that the CCSS should be subjected to examination, trials, and revisions by educational practitioners. In this research, I sought to determine if a student's performance at the state assessment may be predicted from the student's performance on the school assessments. Thus, I used multiple linear regression to explain the association between the predictor variables (a student's midyear and end of the year assessment scores, race, gender, and socioeconomic status) and the criterion variable (observed score on the state assessment). The results of my research indicated a strong and predictive relationship, this may help illustrate how well the current standards are being implemented.

Purpose of the Study

The purpose of this research was to determine how well a student's midyear and end of the year assessment scores, race, gender, and socioeconomic status may predict the student's overall state assessment scores. Four multiple linear regression models were implemented with the midyear and end of year assessment scores, race, gender, and socioeconomic status as predictors, and state domain score as criterion. One model predicted the overall scores while the remaining three models were used to predict the respective scores in geometry, algebra, and number sense. I received district and state assessment data for the one participating South Florida charter school.

There are many standards in education that can be researched. Sforza et al. (2016) focused on reading, writing, and math because these subjects have high stakes tests attached to them. Mathematics is within many job requirements, but many people

struggle with this subject (Nahornick, 2016). Mathematics is important for many careers, yet students are afraid of this subject. Nahornick (2016) stated that although society holds mathematics in high regard, it is not uncommon to hear "I am bad at math" (p. 110). The weight of mathematics in a student's education and career is pertinent to the advancement of this core subject. My research may add to the current data already existing on mathematical assessments. With a strong and significant predictive relationship, students, teachers, schools, and districts may be able to continue what is working as well as copy working procedures to areas where the relationship might lack. Ultimately, results from this study may allow stakeholders to properly prepare to meet the current mathematics standards.

Each standard within the CCSS requires that students understand specific prior mathematical concepts (Louisiana State Department of Education, 2013). Students who are deficient in mathematics will have a difficult time learning new standards because of the lack of prior knowledge. Although this is true for many subjects, it is particularly so for mathematics. The mathematical standards are designed around the progression from grade to grade (CCSS Initiative, n.d.). The students' current standard is built upon foundations built from previous years. Richardson and Eddy (2011) stated the importance of all core subjects, but focused on mathematics because it is an area of lower student achievement in comparison to other school subjects. Their article called for analysis of data for mathematics. Without current data, educational decisions and policies may not be reflective of the current educational landscape. However, with current data, more accurate decisions may be made regarding the mathematics environment. Findings on

whether school assessment scores are predictive of state assessment scores is research that is meaningful and relevant to the current educational stage.

There are two central goals of the CCSS policy (Main, 2012; McCracken, 2014). The first is to set a unified educational standard across the nation and the second to create critical thinkers so that they can compete in the global market. The second goal was not be investigated in my research. The purpose of this research is to provide insight from a middle school's math student scores that may then be used to generalize about other middle school math student scores.

Research Questions and Hypotheses

Based on the problem statement and the purpose of this study, the following are the research questions and their respective hypotheses.

RQ1: Do the observed overall score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination predict the observed overall score on a mathematical state assessment?

$H_0$1: The observed overall score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do not predict the observed overall score on a mathematical state assessment.

H_a1: The observed overall score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status

individually or in linear combination do predict the observed overall score on a mathematical state assessment.

RQ2: Do the observed geometry domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination predict the observed score of the geometry domain on the state assessment?

$H_0$2: The observed geometry domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do not predict the observed geometry domain score on a mathematical state assessment.

H_a2: The observed geometry domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do predict the observed geometry domain score on a mathematical state assessment.

RQ3: Do the observed algebra domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination predict the observed algebra domain score on a mathematical state assessment?

$H_0$3: The observed algebra domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do not predict the observed algebra domain score on a mathematical state assessment.

H_a3: The observed algebra domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do predict the observed algebra domain score on a mathematical state assessment.

RQ4: Do the observed number sense domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination predict the observed score of the number sense domain on the state assessment?

$H_0 4$: The observed number sense domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do not predict the observed number sense domain score on a mathematical state assessment.

H_{a4}: The observed number sense domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do predict the observed number sense domain score on a mathematical state assessment.

Theoretical Framework of the Study

The framework for this dissertation centers on the implementation of standards through curriculum and the testing of these standards. This research focused only on math to generate a better picture of one subject. This choice narrows the perspective implementation of standards allowing focus on key domains geometry, algebra, and number sense.

The U.S. Constitution gives the responsibility of education to the state and district governments (Robbins, 2013). Federal influence began to increase with the influence from educational progressives such as Dewey (Robbins, 2013). Kretchmar (2015) wrote on the philosophy of progressives and their belief that schools should prepare students to participate in society and help create an egalitarian democracy. This philosophy was used to create both a more useful and complete person. CCSS Initiatives (n.d.) referred to their standards as college and career readiness standards. For students to compete for career positions, there need to be standards that can transfer easily across varying locations in the country.

The idea of having the educational transferability gave birth to the national standards put forth by the National Education Department and adopted by many states. Some of the states adopted the standards with only small edits to the standards. Only eight states have yet to adopt the CCSS as indicated by the CCSS Initiative (n.d.) national map.

Based on the information put forth by CCSS Initiative (n.d.), the standards focus on six criteria: (a) are research-and evidence based, (b) are clear, understandable, and consistent, (c) are aligned with college and career expectations, (d) are rigorous and applicable to higher order thinking skills, (e) are based on the current curriculum, and (f) prepare students for success in the global economy and society. Three of these criteria are directly tied to this research. The first criterion is the research and evidence that drive these standards. This indicates the need of constant research and evaluation of the educational process. The second criterion is in the rigor and higher order thinking in

problem solving. The student must demonstrate conceptual (understanding of the topic) and procedural (speed and accuracy of the calculations) knowledge, as well as apply learned (solve real life situations) knowledge. In addition to the basic understanding of a standard, a student must apply procedural knowledge, which is reflected in the speed and accuracy of the calculations. A student must ultimately apply learned knowledge to solve problems in real life situations. The last criterion is that the standards are based on the current curriculum (CCSS Initiative, n.d.). This is important because these standards build on previous topics learned. This means that students will gain knowledge by building upon previous knowledge.

For each topic, there has been a deeper focus on the implementation of critical thinking to the standards. Dewey (1910) used the term *abstract thinking* to refer to deep thought and stated that it is unnatural to have instruction without thought. When a person is using abstract thinking, the person will begin with a nonconcrete idea and seek to make the idea clear, concise, and concrete (Dewey, 1910). Once an idea is concrete, the person has reached understanding of the topic. The opposite can also be true, as familiar topics can be over thought and become strange, unsolved, and back to abstract. Dewey also used the term *engaged intelligence*, which enables an individual to participate effectively with their surroundings (Robbins, 2013). Engaged intelligence allows an individual to learn actively and continually instead of just using memorization to recall concepts. Dewey's learning approaches were based on the principles of constructivism, which are also the foundation of the assessments being analyzed for this research. This foundation can be seen in the focus that has been put forth by the CCSS Initiative (n.d.). Constructivism has

the learner build upon prior experiences, thus creating new knowledge. This is the framework of the standards that I am researching. The constructivist approach to learning is an active process versus the passive procedural-formalist curriculum of past times (Grady et al., 2012).

Nature of the Study

In this study, I used a quantitative approach to study the predictive relationship, if any, between district assessments scores and state assessment scores. I implemented the quantitative nonexperimental, secondary data analysis methodology. Data from the one participating school was received from the South Florida charter school database.

There were two scores that were analyzed. The first was the overall student scores of each assessment. Within each assessment there was also three domain scores analyzed. These domains were algebra, geometry, and number sense. To establish if there is a predictable relationship between student district assessment scores and select demographic variables and state assessment scores, I implemented four multiple linear regression statistical models.

Definitions

The selected South Florida school district mandates that district assessments be given three times a year by Charter Schools USA (CSUSA, n.d.) to students. CSUSA operates 90 charter schools in six states. CSUSA produces the beginning, midyear, and end of the year assessments for elementary, middle, and high schools for these schools (CSUSA, n.d.). In this research, I focused on the midyear and end of year mathematics assessments for 8th graders. I left out the beginning of year assessments because of

summer learning loss. Gershenson and Hayes (2017) spoke about the loss of mathematical knowledge during the summertime in which students are not in schools. Thus, it would be counterproductive to assess a student who has lost knowledge as well as evaluate an assessment that is 10 months away from the state assessment.

The state assessment is administered once a year towards the final months of the school year. The 2018-2019 school year scores were the only year analyzed for this research. Each state creates this type of assessment to measure the mathematical knowledge of individual students. The state of Florida follows suit and creates both the process and the instrument for this assessment. This assessment is then given to local schools to administer to students. In this research, this assessment is referred as the state assessment. The selected school holds the results within the students' records office. Once I received written permission from Walden's Institutional Review Board (IRB), I requested and received the assessment data and student demographics from the students' records office.

The assessment scores in the midyear and end of the year assessments, race, gender, and socioeconomic status are the predictor variables, and the criterion variable is the score obtained on the respective state assessment. Thus, there were four student scores that were analyzed: (a) overall assessment score, (b) algebra, (c) geometry, and (d) number sense scores. Although the races included in the study were White, Black, Hispanic, and other; the variables were dummy coded as White and Non-White. Similarly, gender was dummy coded as male and female. Finally, the socioeconomic status of a student was dummy coded as student paying full or reduced lunch.

Assumptions

There are several assumptions I made for this research. The first was that the district's assessments are given to all students with the same guidelines and specifications. Similarly, I assumed that students are taking each district assessment as seriously as the state assessment. Finally, I also assumed that the scores are not being altered in any manner.

Assumptions must also be made about the state assessment and data collected from the state of Florida. One assumption is that the state assessment is created by a highly reliable and reputable team and/or organization. Similarly, I assumed that the assessment is highly confidential and not public. Finally, I also assumed that the scores provided have not been altered in any manner.

Scope and Delimitations

Data for the study was received from one school chartered by CSUSA in South Florida. The reason for this selection was because I work in this selected South Florida charter school. I am a middle school mathematics teacher. Working here gives me accessibility to data needed for this research. Although I may have been the teacher of several of the students included at some point, I was not their current teacher during the time the data were collected. At that time, I was teaching advanced students which were excluded from this dataset because they did not take the eighth grade assessment. The school principal, see Appendix A, granted me written permission to receive and analyze the student assessment scores and demographics of the school being used in this study.

The National Assessment of Educational Progress (NAEP) gives this current research a guideline to research mathematical assessment scores. The NAEP is a national assessment given by the National Center for Education Statistics (NCES) which is part of the U.S. Department of Education (citation). The NCES randomly selects students across the United States who take the NAEP. With the scores of the NAEP the NCES can compare results from state to state. The state test assesses students but does not provide a comparison with nationwide results. The NAEP math assessments are given to fourth, eighth, and twelfth grades (NCES, n.d.). The reason for the selection of eighth grade data is to have the option to align this research with the NAEP scores later. Following this guide, the data for this research was aggregated from eighth grade students who attended the selected suburban South Florida charter school. The scores came from assessments that the school has already administered. I received existing archival data from the students' records office.

Limitations

The biggest limitation of this research was the data itself. After receiving clearance from the Walden's IRB office, the students' scores and demographics were received from the selected school. Upon my request, the data was given to me to conduct the statistical analyses for my dissertation. This data needed to be sound data. Since I did not create or collect the data myself, this creates high dependency on assessment administrators, students, and school's student records office to act in good faith.

Significance

Educational policies make an impact on students, parents, teachers, school administration, and school districts. Education became a national topic in the United States beginning in the early 20th century when federal control increased (Robbins, 2013). Educational policy is the skeleton in which educational principles have been built upon. This makes research and their outcomes highly required to support or not support educational policies. Main (2012) recommended time for scientifically based research on the standards before using them on high stakes tests. The strength of the prediction and its statistical significance in this study are key requirements in determining and building confidence or acceptance. In other words, if district assessment scores have a strong and statistically significant predictive relationship to the state assessment scores, this research may add confidence in the current process and results for students, teachers, and school administration. The current South Florida district assessments are an educational procedure in need of review so that educational professionals may determine if they are being effective. This research sought to add significant data that the educational community may use.

Thus far, there has been a divided and critical opinion of the common core educational policy (Main, 2012). The opposition to the common core has come from the lack of information and explanation on how the standards would improve education (McGuinn, 2015). This research did not review the specific standards but to determine whether the district assessment scores and state assessment scores have a strong and statistically significant predictive relationship. The results of this research may help the

level of the common core success more clearly to the students, parents, teachers, school administration, and school districts.

There are four reasons why research on the common core policy is timely and significant (Rentner et al., 2014). First, there is a need for examples of what is working to emulate. Local schools will often look at the district and state government for material and resources to carry out the educational policy. Second, policy makers want to assist schools but are not experts in education. This research may assist them in making decisions. Third, research may help shed light on the common core debate with current data. Fourth, the impact of educational research beyond the common core policy. Technology, assessments, and low performing schools are areas in which this research would impact (Rentner et al., 2014). The research in this dissertation included students' mastery, race, gender, and socioeconomic status. This was done to contribute current data to the education field for the use in varying policies and programs. This research may bridge the gap between the educators and policy makers.

The results from this research may potentially benefit several parties, some more directly than others. The results may give current data for lawmakers as well as state and district administration to make decisions regarding assessments and student achievement of standards. The results may directly support teachers in their goals to increase student achievement. The results of this study indicate strong and statistically significant predictive relationships between and among the variables, teachers may also be able to follow these relationships, for student success. The results may also provide data needed for the district and state to make necessary changes. Once the mentioned changes are

made, students who need to master the standards would have the most to gain from the results of this research. Over time, this may yield greater mathematical advancement and overall student mastery. The data provided may also assist parents who want to help their students succeed. I specifically sought to find what is working so it may be continued to be applied or change what may need to be changed. The strong and statistically significant predictive relationship, results of the study may demonstrate what assessment processes may continue. No significant predictive relationships of any of the predictor variables may help indicate what needs to be changed or continued. There is always a need for current results, which results of this research has made available. Results of the study may be of positive social change for students and parents by adding confidence in the mathematical assessments. The results of this study may provide positive social change by empowering teachers and school administration with added support to the current mathematical assessment process. The results may help educators to identify points during the school year of need for extra remediation in mathematics knowledge and skills, and whether these remediations are more relevant to some groups of students. Further, depending on the characteristics and needs of the students, student-centered remediations may be identified and applied.

Summary

As assessments have become an integral part of the U.S. education system. Thus, it is of high importance to determine if the district assessments can predict state assessments. In this study, I attempted to determine if there is a predictive relationship between student's midyear, end of the year district assessment scores, race, gender, and

socioeconomic status, and student's assessment state assessment scores. A multiple linear regression with district midyear and end of year assessment scores, race, gender, and socioeconomic status as predictor variables and state assessment scores as the criterion variable, allowed me to determine the predictive relationship between these variables. This research will add to the body of knowledge for policy makers to make decisions.

In the following chapter, the history and the implementation of the theoretical foundation of this study will be discussed. Chapter 2 will also include extensive relative research that will support the research questions presented.

Chapter 2: Literature Review

Introduction

Since the implementation of the No Child Left behind act (NCLB) and the CCSS, there has been a continuous need to review the impact of assessment on student achievement. The NCLB was initiated in 2001 under the administration of President Bush (Harman et al., 2016). It was the first nation-wide policy to apply high-stakes assessments. In 2010, the CCSS were implemented under the administration of President Obama. The CCSS moved in the same direction as NCLB with little review on the overall outcomes (Harman et al., 2016).

The purpose of this research was to investigate if there is a strong and significant predictive association between district assessment scores and state assessment scores and whether race, gender, and socioeconomic status contribute to predicting the students' state assessment scores. Senator Stephen of Utah accentuated the need for data so policy makers can make informed decisions (Exstrom & Thatcher, 2014). This research can potentially provide data that may support policy makers' decisions regarding student assessments and state standards. With this information, teachers may also use the findings of this study to identify areas in which to assist their students.

In this chapter, I review the literature related to the comparison of student scores on mathematic assessments. The impact of district assessments and state assessments is also discussed. VanDerHeyden et al. (2017) highly recommend including demographics in student assessment research. This current research includes race, gender, and socioeconomic status in the multiple regression models. In my discussion of the

constructivist theoretical foundation, I provide an in-depth review of the beginnings of the constructivist theory. I also review the major views within the constructivist theory and the rationale for its usage in this study.

Literature Search Strategy

I attained the literature for this research through the Walden University Library data bank. I also used Google Scholar. Most articles used for this research are from the year 2014 and newer. The articles that are from prior years were included in this research to show the original source of the idea being presented. I searched all sources that were peer reviewed.

To find these articles, the following key terms were placed into the search engines: *constructivist theory education formative assessment, No Child Left Behind assessments, mathematics, Piaget, Vygostky, constructivism, formative assessment, summative assessment, No Child Left Behind, Common Core State Standards, mathematical assessments, USA,* and *demographics.*

Theoretical Foundation

There are numerous theories in education. I chose the constructivist theory because it is the theory in which the district assessments and state assessments have been formed. These assessments are directly linked to the implementation of the CCSS. The CCSS Initiatives standards are intended to build college and career readiness standards (CCSS Initiatives, n.d.). This philosophy is being implemented to create complete academic individuals. Flores-Koulish and Smith-D'Arezz (2016) cited the usage of the constructivist teaching methods, such as critical thinking, as the expectations of the

CCSS. There is a close relationship between assessments and pedagogy (James, 2006). Harkness' (2016) research demonstrated that classrooms that implement the constructivist approach attain higher achievement versus classrooms that do not implement a constructivist approach. More specifically, students who were in a constructivist environment had significantly higher scores on standardized assessments in comparison to their peers without a constructivist environment (Harkness, 2016). Furthermore, because the standardized assessments were created with a constructivist foundation, a teacher who provides a constructivist environment will potentially help their students achieve higher academic scores in district and state assessments. This is the reason for choosing the constructivist theory as a basis for this research.

The origin of constructivism is extensive. The topic has appeared in different times and manners within formal and informal education environments. The constructivist theory dates to Socrates in ancient Greece (Harkness, 2016). Socrates implemented an innovated approach to asking questions that would lead to complex ideas. The teacher is not to give answers, but rather, pose questions. The teachers might not even know the answer, but they are expected to help the student explore the possibilities and uncover knowledge. Houseworth (2015) argued that the constructivist approach is a collaborative one. There should be conversations that bridge prior concepts to applications of knowledge. In this manner complex problems can be more easily solved by a group working together. Having a student-centered classroom, where the teacher facilitates rather than leads, is the current application of constructivism (Houseworth, 2015).

An influential philosophy in education reform came from the American philosopher, psychologist, and educational reformer, Dewey. Dewey's theory came about in the early 1900s when the federal government dramatically increased their influence in the nation's education. Hornbeck (2017) detailed the evolution of education within the United States. Early in the country's history, education was managed within the local towns. By the 1800s, individual states had adopted public education. It was not until 1865, after the Civil War, that the Federal Department of Education was created. In 1980, the Department of Education became a presidential cabinet position.

As the federal government began to increase its influence on the nation's education, Dewey had a big impact on the nation's educational philosophy. Dewey's core belief was nurturing each student's interest and talents while preparing a well-adjusted civic-minded individual (Tampio, 2017). In this philosophy, schools should not only educate students to pursue their own path, but also be social community minded. Kretchmar (2015) wrote those with the belief that the education system should prepare students to partake in all societal areas, including government, businesses, and civic duties. This action by individuals would promote an egalitarian democracy. To do this, the classroom must be focused on the student. The intent of using this philosophy in the classroom would be to produce well-rounded individuals. According to Tampio (2017), the student is the sun around which education should revolve. Instead of simply presenting the information, the teacher should discover and cultivate the student's capabilities. To apply Dewey's philosophy, critical thinking and real-world problem skills must be taught (Houseworth, 2015). Teachers under the CCSS devote more time to

critical thinking versus pedagogy prior to the CCSS (Nichols, 2017). The current curriculum and standards have in their foundation the constructivist paradigm (Houseworth, 2015). Moreover, Dewey's student-centered educational philosophy was used to shape what has become the current U.S. education system philosophy.

Constructivism breaks down into two categories (Bitter, 2018). The first is with Piaget's cognitive constructivism and the second with Vygotsky's social constructivism. The constructivist theory was heavily influenced by Piaget (Bozkurt, 2017). Piaget believed that knowledge was not found, but made. Under this philosophy, knowledge is the product of a person's own cognitive reasoning (Bozkurt, 2017). With this core belief, constructivists assert that knowledge is not passive, but built by focusing on a subject (e.g., mathematics). There is no room for teaching, but individual, independent work and experiments (Bozkurt, 2017). The cognitive constructivism theory poses that the mind obtains new information from existing knowledge. Within social constructivism, the priority is given to the dialogue. In other words, the interaction between participants is responsible for reaching new knowledge (Bitter, 2018). Under Piaget, cognitive constructivism is also called individual cognition.

Vygotsky had a different approach to constructivism (Bozkurt, 2017). Vygotsky believed in social constructivism. In this philosophy, social interaction and guidance from a more skilled peer or teacher is given priority. Through this social interaction, the individual will gain intellectual development. The zone of proximal development describes the development level based on the guidance of a more knowledgeable person (Bozkurt, 2017). Within the social constructivism, every development is first seen on a

social level and then on an individual level (Bozkurt, 2017). In other words, for students to replicate higher functions, they must first see it being applied in their surroundings. Although cognitive and social constructivism seem to oppose each other, they both require prior knowledge to build new knowledge. The main difference between the philosophies is the source from which the knowledge is acquired. While cognitive constructivism has knowledge streaming from within, social constructivism has knowledge coming from the student's surroundings.

Major Theoretical Propositions

The constructivist learning theory is grounded in the principle that the student will learn when they construct new knowledge by using prior knowledge and critical thinking (Reel, 2010). Teachers using the constructivist learning theory will identify that students will learn the curriculum in a manner that is related to their own experiences and goals. The teacher's knowledge of this learning theory might be consciously or unconsciously highly impacting the achievement level of their students.

The current application of the constructivist theory is the way students acquire knowledge and how the teacher will provide knowledge (Bitter, 2018). Critical thinking will happen more effective as a group than as an individual. Although there are two opposing views of constructivism with Piaget's cognitive constructivism and Vygotsky's social constructivism, both are essential (Bozkurt, 2017). Having internal learning is necessary but having a guide is pivotal in learning. Scaffolding is where a teacher enables students to carry out a task with a gradual decrease of guidance. The use of scaffolding has become an effective tool for teachers to use (Bozkurt, 2017).

The CCSS has made it known to the states and districts that the basis of the standards is the critical thinking. CCSS Initiatives (n.d.) referred to their standards as college and career readiness standards. The standards used for the district and state assessments follow the CCSS, which are made to prepare individuals to use critical thinking and participate in all types of societal matters.

There have been many opponents of the CCSS. Piaget's cognitive development stages are the principles of the CCSS (Nichols, 2017). Nichols (2017) argued that in Piaget's philosophy, a teacher must wait for a student to be ready to learn a topic. This, however, is at odds with the CCSS that has students nationwide prepare for the same standard within the same timeframe. Nichols suggested that the thinking should not be by year but prekindergarten to Grade 12 as a collective.

Bozkurt (2017) reviewed the constructivism through the lens of the current mathematical paradigm. Bozkurt presented cognitive and social ways in which the constructivism is being applied in classrooms and concluded that interactive ways of learning is the best support of individual learning. In other words, both approaches working together will yield the best outcome.

Deane et al. (2015) established a direct link between learning models and assessments. In their research they presented the ways that key practices are applied in classrooms. They demonstrated the association between student practices and learning progressions and finally assessments. Deane et al. also stated that making the connection between instructional practice and assessment is difficult, but this connection will help

teachers enhance student's educational achievement. It was concluded that class practices are key and have a significant effect in assessments (Deane et al., 2015).

Literature Review Related to Key Variables

In this section I present peer reviewed literature that will further the understanding of student achievement on mathematical assessments. Each section was selected because of the impact and significance that each section contributes to this research. The following literature will be used as a guide to this research. In turn, this new research will add to the collective data so it may be used for future research and policy decisions.

Mathematical Standards

The mathematical standards that are now in place have been set by the CCSS. Cipriani (2015) presented a detailed history of what are now the national math standards. The first time the idea of national standards was presented was in 1980 at a National Council of Teachers of Mathematics conference (Cipriani, 2015). In that same conference it was proposed that mathematics should shift from computation to problem-solving. Several projects were implemented after this such as the Algebra project which began to teach seventh and eighth graders algebra. As a result of this 39% of the students in this project were placed in Algebra Honors.

In 1991 the National Council of Teachers of Mathematics published Professional Standards for Teaching (Cipriani, 2015). In this publication there were examples of how to teach mathematical concepts in new ways. The response to this publication was immense. Teachers had questions and wanted to know more about the methods proposed. In response to these questions the National Council of Teachers of Mathematics

published The Assessment Standards for Mathematics. In this second publication new ways to assess were suggested. The teacher was to ask questions, listen to the response, use multiple sources of assessments, and use real-world problems. The National Council of Teachers of Mathematics reasoned that the assessment must match standards and teaching models (Cipriani, 2015).

The biggest change came in 2009 in a collaboration between the National Governors Association and the Council of Chief State School Officers. The goal of this collaboration was to create a framework to prepare students for college and work as well as have consistent standards nationwide. In 2010, the CCSS were released giving teachers' mathematical standards by grade broken down into domains. The domains set are algebra, geometry, and number sense for eighth grade.

The purpose of these standards is to establish more rigorous education for students (Lee, 2016). The CCSS standards are meant to prepare students for higher education and to compete with worldwide counterparts. The standards were created to be vertical for a long-term vision (Lee, 2016). This means that the standards span across many grade levels. This was done to allow teachers time to prepare students for higher education and the workforce. With this goal in mind every year's mathematical learning is important because it builds on previous mathematical concepts. Lee (2016) argued the importance of elementary and middle school years to the success of the overall mathematical standards. These early years will be the foundation for the standards that follow.

Groß et al. (2016) conducted research on the diagnosis models within Austrian baseline math tests. Moreover, the researchers presented a noncompensatory deterministic input, noisy 'and' gate (DINA) model which falls under the cognitive diagnosis model. Their research reviewed two types of assessment models used. The first is item response theory which, when applied to an assessment, focuses on one item at a time as opposed to the overall student score. The second method reviewed was the cognitive diagnosis model. This model assigns the students an overall score; but, unlike a percentile, the score is a profile which will place each assessment taker into a group. Everyone in the group will have similar educational achievements. The cognitive diagnosis model will specify the skills and attributes based on what was required to solve the problem (Groß et al., 2016). In this manner the score was used to demonstrate the weaknesses and strengths of the assessment taker within the group. The DINA model applied both two parts to each question. The Austrian baseline assessment follows the Rasch model of conformance (Groß et al., 2016). This is where assessment items are categorized into groups. The Austrian baseline math assessment is broken down into the following four domains: numbers and measures, variables and function dependencies, geometry, and statistics. The research put forth by Groß et al. (2016) highly supports the domains within the research questions of this current research. The standards for both assessments that I researched were broken down into domains as well. To find predictive association in assessment scores, I compared data in the following math domains: algebra, geometry, and number sense for eighth grade. The research set forth by Groß et

al. (2016) was made up of eighth grade student assessment scores. In my research, I also used data from eighth grade students.

Within the Austrian baseline assessments, there are the following subcategories of model building, calculation, interpretation, and augmentation (Groß et al., 2016). The CCSS has provided subcategories as well. The following are the subcategories that are embedded in the standards: make sense of problems and persevere in solving them, reason abstractly and quantitatively, construct viable arguments and critique the reasoning of others, model with mathematics, use appropriate tools strategically, attend to precision, look for and make use of structure and, look for and express regularity in repeated reasoning (Akkus, 2016).

There are several pros and cons of having the math CCSS (Akkus, 2016). The first pro is the broader content available coming from 41 states that have implemented the math standards. Since material is applicable to wide audience, educators can share resources with different schools, districts and states. Having more options leads to higher quality content and assessments. It also promotes collaboration in two ways. First, it allows the collaboration throughout the large network created. Second, since the standards promote a progression for the mathematical concepts, teachers may collaborate up and down grade levels.

Akkus (2016) also presented a survey in which 12,000 mathematics teachers participated. Over 90% said they were in favor of the CCSS. This percent of teachers stated that having the standards kept expectations clear and consistent in what students should learn. Another pro would be that the new standards have less topics a year

(Akkus, 2016). This allows teachers to stay on the same topic longer. Another pro is that the CCSS still gives local autonomy to schools and teachers. Since the CCSS does not give any specific materials, acquiring materials is left up to the local educators. Although this has been troublesome, it gives educators the power of input and contribution to their students' learning. Although the article by Banks, LaFors, and Education Trust-West (2015) is a non-peer reviewed article, the direct classroom research is valuable. The CCSS encourages thoughtful questioning and collaborative questioning. The CCSS brings the same standards to everyone that were previously available only to upper income students. With the new expectation, students in lower income demographics are given the same level of standards as everyone else. Although this can be a challenge, it also provides the opportunity to achieve the high standards.

As with any policy, there have been several cons that have come in the process of applying CCSS. In the survey that Akkus (2016) presented, teachers identified lack of support in teaching the CCSS. Teachers also said that it was a challenge to acquire curriculum material. In the same survey out of 12,000 teachers, less than half felt prepared to teach the CCSS. Teachers have been given confusing guidance by textbooks (Akkus, 2016). Publishing companies have different ideas on how to present the material. Educators must then select from these options. With little or no preparations, the best decisions to choose materials cannot be made. Without real guidance, this process has been mostly trial and error at the expense of wasted learning time. Akkus (2016) recommended that authorities give more focus on the math CCSS materials that will be used. Akkus (2016) also recommends that teachers be given time to reflect and converse

on the progress of the math CCSS. In the non-peer reviewed article put forth by Banks et al. (2015), the authors critiqued the math CCSS stating that teachers have been challenged to teach new standards to students who have not had these standards applied until now. This has caused gaps in student understanding and overall achievement. In this same article the authors illustrate the issue of textbooks not having the correct alignment. Publishers will state that their books are following the CCSS, but it is up to the teachers once again to separate what is useful and not useful to achieve mastery of the standards.

Types of Assessments

In modern education, formative and summative assessments have become the tools that will indicate student achievement. The district assessments and the state assessments that are being researched in this study are both formative and summative assessments. An explanation of both types of assessments will be provided. The midyear district assessment is both a formative and summative assessment. The end of year district assessment and the state assessments are both summative assessments.

Formative assessments are tools that provide feedback so teachers can modify their teaching to the student needs (Andersson, & Palm, 2017). These authors further stated that there are two manners in which formative assessments are applied. The first is using tests with questions that gather what a student knows. This will then allow teachers to provide feedback. In this manner there is high importance in the assessment and the collection and interpretation of the data. The second manner to which apply a formative assessment is by varying evidence that will gauge a student's understanding. In this manner a faster modification to the learners thinking can be given. The first format is the

way the midyear district assessment is being implemented. This gives high priority to the interpretation and usage of the data in order to facilitate feedback. CCSS based lessons have given way to better formative assessments (Ateh & Wyngowski, 2015). The researchers argued that in the process of adjusting instruction, teachers who implement formative assessments are the most effective. The CCSS has built in the need for formative assessments leading to the end of a unit or course.

Summative assessments are evaluations of performance with the purpose to assign students a score based on their knowledge of content (Schoenfeld, 2015). These types of assessments are at the end of a unit or a course. Summative assessments provide first perspectives into student achievement (Marinho et al., 2017). In summative assessments students must respond to answers, subsequently teachers will give a final score to the student. The score will not only be for the assessment but representative of the unit or entire course. Examples of summative assessments are unit test, a class final, SATs. The state assessments are summative assessments. These assessments are a form of accountability from the teacher to the student but also from administration to the teacher. Marinho et al. (2017) critiqued summative assessments for the stress that they produce to students and teachers.

With the introduction of the No Child Left Behind (NCLB) policy in 2001 came the introduction of federal influenced state assessments. Prior to the NCLB policy, states were free to follow their own curriculum and create their own assessments (Schoenfeld, 2015). Under the NCLB, states followed the federal guidelines in curriculum and assessments in order to receive federal funding. Not only were the student promotions to

the next grade based on state assessments, but also teachers and administrators' salaries began to be dependent on the scores as well. This changed the landscape of education with many teachers and schools teaching to the test. The assessments became known as high stakes tests because of the high impact the assessment results had on all stakeholders. Markowitz (2018) presents both the negative and positive impact that the NCLB has had on student achievement. On one side, this paradigm has led to a narrower curriculum with more time spent on learning for the test. Teachers have less time to know their students and engage them holistically. On the other side, the increase on academics has created higher student achievement in the areas of focus.

When the CCSS was adapted, it continued with what was being implemented by NCLB regarding assessments. Within the CCSS, the federal government set up the initiative called race to the top. This was designed for states to compete for federal money based on student performance on state assessments. The CCSS implements high stakes assessments and that having high stakes assessments are a big factor on how the standards are implemented and taught (McDuffie et al., 2015). The curriculum did narrow because of teaching to the test but also stated that there was expansion of curriculum in some cases (McDuffie et al., 2015). Another challenge of the CCSS was the short time between policy development and classroom application. Teachers were presented the material and expected to teach it at the same time. McDuffie et al. (2015) gave a comparison of Japan in the same scenario giving their teachers 3 to 4 years to prepare and critique the standards. This lack of time has given critics the ability to state that the CCSS are unproven and more research is needed. McDuffie et al. (2015)

concluded that teachers found the CCSS much more rigorous than prior standards. The biggest grievance from teachers is the lack of support and aligned materials needed to teach the standards.

Mathematics Assessments

Mathematics is a core subject in the United States education; therefore, mathematical assessments have a big impact on individual students as well as schools and districts. Daro and Burkhardt (2012) gave a description of each mathematical domain being assessed in the CCSS. The following Content and Representation of Mathematical Domains definitions have been provided by Cpalms. (n.d.). It states number sense domain containing number concepts, representations relationships, operations, computation and estimation; Geometry domain as containing shapes, properties of shapes, relationships, spatial representation, location and movement, transformation and symmetry, visualization, spatial reasoning, and modeling; algebra domain contains patterns, relations, functions, ratios, proportions.

There are test theories that are used to construct mathematical assessments. The classical test theory and the item response theory are two theories that are widely used in assessments (Choi, Lee, & Park, 2015). Both theories are also used in the construct of the assessments within this research but the latter one is used more. The classical test theory states that a student's score is the addition of the true score, the assessment with no construct errors, plus the assessment errors. This theory has the assumption that all assessments have errors within the construct. The item response theory focuses on the knowledge of the question item versus the whole assessment score. Concentrating on one

item allows for more data to research. Both the district and state assessments break down the standards tested into domains which allow for a more accurate representation of student achievement within the standards tested. Choi, Lee, and Park (2015) critique both the classical test theory and the item response theory stating that there is a need to further examine mastery using more precise tools.

Traynor (2017) conducted research on differences in state curricular content standards and student assessment scores. Traynor (2017) defines the curricular content as the standards that legislators intended for students achieve. The author argues that when an assessment is delivered over a great geographical area there may be differences not in the content itself but in the student's opportunity to learn. The author posed the question whether instructional sensitivity and curricular validity make a significant difference in student assessment scores. The National Assessment of Educational Progress (NAEP) and the Third International Mathematics and Science Study (TIMSS) were used for this research. The items were divided in the following groups; Number sense, Properties, and Operational Measurement; Measurement; Geometry and Spatial Sense; Data analysis, Statistics and Probability; and Algebra and Functions. The complexity of items was split into low, moderate, or high cognitive complexity. The results of the research indicated that there was little evidence that curricular content made a significant effect on assessment scores. Traynor (2017) advised that more research is needed to assess the opportunity to learn by students taking state assessments because of varying factors. These factors may include classroom instruction, individual cognition as well as states changing the construct of their assessments.

Shivraj (2017) put forth an article that aligns the CCSS in mathematics to an international assessment called Programme for International Student Assessment (PISA). Since the year 2000 students from different participating countries have taken this math assessment. This has been done to rank student assessment scores by country. The results from 2012 had the United States ranked at 26 out of 34 participating countries. The mathematical domains being assessed are number sense, geometry, algebra, probability, and data analysis. The United States is far below countries like China, Singapore, Canada, Australia, and Japan (Shivraj, 2017). The purpose of their research was to find what is lacking from the CCSS assessments and curriculum that will prepare students to the real world. The researcher concluded that the standards and implementation of the CCSS was not aligned to the PISA. The PISA not only assesses the domains given but also how well students can use their knowledge in order to solve unfamiliar settings. The CCSS does provide instruction in this area but it takes time to see the outcome of this instruction. If the United States is to move up in ranking, there is a need for CCSS mathematical standards and assessments to be further addressed (Shivraj, 2017).

Anselmo et al. (2017) conducted relevant investigation to this current research. Their research conducted several types of Curriculum Based Measurements (CBM), which are short assessment, that measure a specific skill. Their research examined the predictive validity of the benchmark scores of the Mathematics Curriculum Based Measurement (M-CBM), Math Concepts and Applications (M-CAP) and a reading comprehension assessment by the company AIMSweb called MAZE to state testing within students in grades 7 and 8. The M-CBM examines how well students produce

accurate answers. The M-CAP examines how well students do on multi-step math problems. MAZE is a reading comprehension that indicates students surface level text comprehension. In previous primary grade research, the MAZE had predictive validity in student mathematical achievement (Anselmo et al., 2017). All three assessments are given in class and take less than 15 minutes. The researchers examined the results of 298 participants broken down by gender and ethnicity. Their results indicate that there was a significant relationship between the M-CMB and the state test. The relationship was weak in both years tested (2012; NC-EOG-M: $r = .21$; 2013 NC-EOG-M: $r = .21$). The percentage of explained variance was 4% in 2012 and 4.5% in 2013. There was a stronger relationship between the M-CAP and the state test. The percentage of explained variance was 42.8% in 2012 and 43.6% in 2013 with similar predictive validity in both years ($r = .65$; $r = .66$). The MAZE scores also had significant but weak correlation to the state test. The percentage of explained variance was of 8.9% in 2012 and 11.6% in 2013. Anselmo et al. (2017) noted that there is extensive research for reading curriculum-based measurement but very little in mathematics. Even with their research, more research is needed to add to the mathematics literature. The comparison and discovering the predictive power of district midyear and end of year assessments to the state assessment may add to the mathematics literature.

Demographics as Factors

Race, gender, and socioeconomic status were used as dummy coded variables in this research. It is fitting that I review these student demographics in prior mathematical assessments research. Gottfried (2016) conducted research to investigate the role of Real

Life Mathematical Instruction (RLMI) on student end of year achievement as well as differences in achievement in regards to gender, race and socioeconomic status lines. Race was categorized by White, Black, Hispanic, and Asian. Approximately $n = 8500$ kindergarten students were selected and followed through fifth grade for data analysis. The frequency of RLMI was annotated as never, once per month, less than weekly, 1-2 times a week, 3-4 times a week, and daily. In fifth grade the students were given a fall and spring mathematics assessments which measured conceptual knowledge, procedural knowledge and problem solving. The topics included number sense, properties and operations; measurement; geometry and spatial sense; data analysis, statistics and probability; and patterns, algebra and functions. Gottfried (2016) indicated that all frequency indicators with RLMI had a significant difference in the spring assessment. Both genders show a benefit from RLMI but males had higher gains in scores. Regarding race, Whites and Blacks had significant difference with Blacks having the highest gains in scores. Hispanics and Asians had no significant difference.

Jian-Hua Liang et al. (2018) investigated the predictive power of cognitive and noncognitive variables to predict performance of eight graders on Algebra state assessment. The cognitive variables were students' seventh grade English Language Arts and Mathematics California state test scores as well as the mathematics sub-scores. The non-cognitive variables were the student's demographics gender, ethnicity, parent education level, participation in the National School Lunch Program (NSLP), special education programs, the Gifted and Talented Education (GATE), English learner (EL), and Reclassified-Fluent English Proficient (R-FEP). The study began with 209,364

students but was brought down to 34,000 because of missing or invalid data. The results indicated that the seventh grade state scores had the most significance to the algebra scores, it accounted for 61% of explained variance.

The sub-scores of number sense, rational numbers and geometry all had significant mean differences in the algebra scores. The rational numbers category accounted for 48% of explained variance on the algebra scores. Jian-Hua Liang et al. (2018) argue that, within rational numbers, students must be able to manipulate numbers, which is the basis in algebra. The second highest predictor was quantitative relationships with 8% explained variance. The third predictor was measurement and geometry with 4% of explained variance. The demographics that had positive coefficients were the following: Asian, GATE, Parent education level, EL, R-REP, gender, and special education. The demographics that had negative coefficients were NSLP, Hispanic/Latino, African American, and American Indian. Jian-Hua Liang et al. (2018) noted that the negative relationships of these variables are consistent with the California Department of Education study of the achievement gap. The gap has been narrowed but continues to be a factor. Jian-Hua Liang et al. (2018) concluded that focus on rational numbers and quantitative relationships will yield better results for students preparing for algebra. It would also be beneficial to focus on these categories for students who are struggling within the Algebra course.

VanDerHeyden et al. (2017) also did research pertaining to demographics. The purpose of their research was to find the most effective and least costly manner of predicting year end mathematics scores. The research was conducted in classrooms from

grades 3 to 5 in urban locations. The students were given a Mathematics Computation (M-COMP) and a Mathematics Concepts and Applications (M-CAP). Both of these assessments are Curriculum Based Measurements. As previously stated, a CBM is a short assessment which measures a specific skill. The researchers included prior year assessment scores as well as demographics. Prior research has concluded that demographics accounted for 75% of explained variance when predicting for mathematics assessment scores in third graders (VanDerHeyden et al., 2017). It was also noted that demographics of students varied greatly. VanDerHeyden et al. (2017) noted the importance of sociodemographic variables as risk factors. VanDerHeyden et al. (2017) accounted for demographics risk factors that were available for their research. They recognized that demographics such as ethnic diversity may yield more accurate results and necessary for this type of research. The researchers included race, gender, English as a second language, and learning disability as demographics in their own research. In this same research there were three screenings given in fall, winter, and spring, respectively (VanDerHeyden et al. 2017). This is compatible to the district assessments of my current research. In the research put forth by VanDerHeyden et al. (2017) there were 182 students with results that indicated that the preceding year's score was the most comparable and lowest cost to predict the current year test score. The winter assessments were second best at predicting scores. This research gives credence in the selection of using only the midyear and end of year in my current research. It was discovered that demographics did not have a significant relationship in their regression models. VanDerHeyden et al. (2017) argued that predictive value of demographics could have

been affected by sample size as well as diversity constraints in their own research. In response to this statement my current research included select demographics.

Bohrnstedt et al. (2015) emphasized the need to research the Black-White achievement gap in math. Although this current research did not focus on this issue, race was included in the study. The current research may add to the literature pertaining to race as factor. Bohrnstedt et al. (2015) conducted an extensive study on math assessments scores and the impact of race. The researchers argued that although this problem has been studied, very little has been done in relationship to school composition. The data used for this research was the National Assessment of Education Progress (NAEP) as well as the Common Core of Data. On average White students attend schools with 9% Black students while on average Black students attend schools with 48% Black students (Bohrnstedt et al., 2015). This means that Black students' families cluster in areas. The achievement in areas with high density of Black students had lower assessment scores but the achievement gap was same between races as in other schools. Regarding gender, Black males had the largest gap in highest density schools. The researchers state that demographic research is important because concerns of re-segregation. The continued study of race as a variable will help understand how it is a factor in student achievement.

Summary and Conclusions

This research investigated the relationship between district assessment scores and state assessment scores and weather race, gender, and socioeconomic status contribute to predicting the students' state assessment scores. This chapter provided a detailed review of relevant topics and bearing research to the current research.

The literature of the constructivist theoretical foundation, which was used to build the assessments, was reviewed. Deane et al. (2015) established a link between learning models and assessments. The researchers demonstrated the association between student practices and learning progressions and finally assessments. The impact of Piaget's cognitive constructivism and of Vygotsky's social constructivism on assessments was reviewed.

The history and content of the mathematical standards that are now in place by the CCSS were reviewed. Types of assessments such as formative and summative assessments were presented as well. This review leads to the path of the current high stakes assessments that are state assessments.

Associated literature to the comparison of student scores on mathematic assessments was extensively reviewed. Various studies on student mathematical assessments and their scores were studied. Research that is closely related to the current research was presented. Traynor (2017) conducted research on differences in state curricular content standards and student assessment scores. Anselmo et al. (2017) research was reviewed in this chapter for their relevant study to this current research. Their research examined the predictive validity of their specific benchmark scores.

This current research adds to the mathematics assessment literature. This research investigated the predictive validity of midyear and end of year assessments to the state assessments. In this manner it has added to the literature pertaining to mathematical assessments. It also added to the literature gap the factor of demographics on mathematical assessments.

In the next chapter, a section on the research methodology is presented. Next are the research questions that guided this study, and a description of participants and instruments used in the study. Data sources and data collection are also discussed. The last section of the chapter is a description of the data analysis plan.

Chapter 3: Research Method

Introduction

The purpose of this current study was to determine if there is a predictive relationship between student's performance on a given midyear district assessment, end of year district mathematics assessment, race, gender, and socioeconomic status, and the state assessment. This research may help determine the predictive power of student's midyear and end of the year assessment scores given the detailed variables. In this chapter, the following categories are explicated: research design and rationale, variables and methodology, population, sampling, procedures for the sampling, data analysis plan, and threats to validity.

Research Design and Rationale

This research includes two main predictor variables (midyear and end of the year district assessments) and three demographic predictors (race, gender, and socioeconomic status). The criterion variable is the student's score on the state assessment. For both types of assessments, the overall scores are used. The individual domain scores of geometry, algebra, and number sense are used as well. All district scores (midyear and end of the year) as well as all state assessment scores (overall and individual domain) are continuous. The demographic predictors in this research are race, gender, and socioeconomic status are categorical, but were be dummy coded.

The research design for this study is quantitative in nature since I am trying to determine the predictive relationship between student's performance on a midyear and end of the year district assessment, race, gender, and socioeconomic status, and the state

assessment. After examining the research questions, the variables, and the relevant literature, a multiple linear regression analysis was chosen as the best type of analysis for this research. An example of using this type of methodology is the research by VanDerHeyden et al. (2017) on student assessments scores, which also used a multiple linear regression model. According to Frankfort-Nachmias and Nachmias (2008), this model is used when the research has two or more predictive variables and a continuous criterion variable. This analysis model allows the researcher to assess the predictive relationship of the predictive variable on the criterion variable. If a relationship is found, the multiple linear regression model may help the researcher discover the predictability of the individual predictive variables as well as a combination of all predictive variables on the criterion variable.

Mathematics standards continue to change state achievement programs and schools need to continuously monitor assessments that measure student's achievement; therefore, these mathematics assessments need to be rigorously examined (Traynor, 2017). Anselmo et al. (2017) investigated the predictive validity in curriculum-based probes for the state high-stakes test. Their results indicated that reasoning was a strong significant predictor while computation was not. Anselmo et al. concluded that more research is needed to determine better predictive validity of assessments, which they call "curriculum based measurements." Acquiring the predictive power of the district's midyear and end of year assessments to the state assessment adds to the eighth grade mathematics assessments and may increase student achievement. Bohrnstedt et al. (2015) noted the importance of demographics in student assessment research and the impact that

this type of research may have on a community. In their research it was stated that current data would help understand and even close the White-Black student achievement gap. It is suitable to add more literature by contributing findings on the impact of student demographics on math assessments. The results of the study may advance knowledge in the mathematical assessment discipline by contributing data and statistics that may be used to advance future educational policy.

Methodology

This section describes the population from which the sample was obtained, as well as the sampling steps taken to acquire the data. This section also includes the data analysis plan.

Population

The U.S. Department of Education administers the NAEP. The math assessment is administered to fourth (primary), eighth (middle), and twelfth (secondary) grades (NCES, n.d.). This research used archival data from eighth grade students to follow previous research, such as that of Grady et al. (2012), NCES (n.d.), Cipriani (2015), and et al. (2016), all of which indicated that eighth grade was the grade in which to conduct mathematical research. The eighth-grade data has the unique position of being the middle set, this allows the data to be significant to grades below and above. The NCES is the only reference mentioned that conducted its own assessment. Grady et al. (2012), Cipriani (2015), and Groß et al. (2016) all used archival data. My study's population was based on the available data for eighth grade students. Students selected from a South Florida charter school, which is part of a network of 56 schools that form its own district

in Florida. The school, which has approximately 250 eighth grade students, is in a middle-class suburb where 54% of the students are minority students and 39% receive free or reduced-price lunch.

Sampling and Sampling Procedures

The selected sample came from a population of 250 eighth grade students. Students who completed the 2018-2019 midyear, end of year, and state math assessment were selected. There was an exclusion of 25 students for whom I was their teacher when they took these assessments. These exclusions bring the number to 225 available students.

To find the required minimum sample size, I conducted a statistical power analysis. If the sample size is too small, generalizability becomes an issue; if the size is larger than needed, resources and time may be wasted (Field, 2013). The required minimum sample size was calculated using the software G*Power 3.1.3. program. The test family implemented was a multiple linear regression, fixed model, R^2 increase, and an F test. Level of significance was set, *a priori*, at $\alpha = 0.05$, and power $(1 - \beta)$ at 0.8, following Field (2013). Using these recommendations, the parameters entered into the software G*Power for this study were the α error probability of .05, and the power $(1 - \beta)$ of 0.8. Gibson (2013) used a multiple linear regression analysis to predict the scores of the California Critical Thinking test with variables such as parent education, family income, and extracurricular activities. They used .15 as the effect size to calculate a minimum sample size. Green-Davis and Sha-Rhonda (2017) also used a multiple linear regression model in their research with the effect size of .15 to find their sample size.

Following prior researchers who have used a multiple linear regression analysis, an effect size of .15 was used in this research to calculate the required minimum sample size. There are two tested predictors and three demographic predictors in my research study, which were also entered into G*Power. With the predictor variables, one criterion variable, an effect size of .15, a power level of .8 and an alpha of .05, the minimum sample size for this research is 68 students. This number represents the minimum required number of participants. However, to err on the side of caution, I randomly selected 100 participants, from the population of 225 available students.

Data Collection

This study used archival data from a South Florida Charter school. The school staff have administered the district assessments and the state assessment. The student data are then stored with the school's student records office. The student demographics are also stored in the student's records office. I work at this school and was given written permission from the principal to access the data needed, through the person in charge . Once Walden IRB granted me permission, I requested the data from the student records office and received it as Excel files.

Data Analysis Plan

Once the existing archival data were received, the variables were entered into the software Statistical Program for Social Sciences (SPSS). Prior to running the multiple linear regression analysis, the data was tested to see if it met the multiple linear regression assumptions. The first assumption was that only relevant variables are included in the study. According to the literature review, the variables selected for this

study are relevant as stated in previous studies. Another assumption was the linear relationships between continuous variables. Linearity was checked via correlation tables that are included in multiple regression models, Pearson's *r* correlation as well as scatterplots. The third assumption was that all variables are normally distributed. It was checked by plotting residual values on histogram chart. To be normally distributed the histogram must approximate a normal curve. The last assumption was homoscedasticity, which is the homogeneity of variance. The violation of homoscedasticity was checked by scatterplot of standardized predicted value by standard residuals (Field, 2013). The assumption is met when the scatterplot indicates an even and random distribution.

When multiple linear regression is used, there is the threat of multicollinearity. According to Field (2013), collinearity occurs when there is a strong correlation between two or more predictors. If this occurs, then the beta weights will not be statistically significant regarding the criterion variable. In other words, I would not know which predictor variable is important. To check for this, a Pearson's correlation matrix was conducted using the predictor variables of this study. The variance inflation factor (VIF) was also analyzed. According to Field a correlation above 5 is cause for concern while a VIF of 10 should be the maximum. A tolerance below 0.2 may be a potential problem and below 0.1 indicates a serious problem (Field, 2013). If the findings encounter a strong correlation between predictive variables, any predictor variable that is presenting a multicollinearity issue would need to be excluded.

Four multiple linear regression analysis models were conducted. The first model analyzed predictive relationships of the overall assessment scores. The subsequent

models analyzed the predictive relationships of the geometry, algebra, and number sense domain scores. Each model had the demographic predictors of race, gender, and socioeconomic status for each student. The races included in the study were White, Black, Hispanic, and other; however, the variables were dummy coded as White and Non-White. Gender was dummy coded as male and female. Finally, the socioeconomic status of a student was dummy coded as student paying full or reduced lunch. The *t*-value and *p*-value in the coefficients table were analyzed to determine which predictor variable, if any, had a significant predictive relationship with the criterion variable.

A multiple linear regression was conducted with the following equation: $y = a + b_1X_1 + b_2X_2 + \ldots b_pX_p + e$. The criterion variables, state assessment scores, are coded (Y). The constant is represented by a. The included predictor variables are represented by X's as follows: midyear assessment scores (X_1), end of year assessment scores (X_2), race (X_3), gender (X_4), socioeconomic status (X_5). The midyear, end of year and state assessments are continuous numerical variables. Race, gender, and socioeconomic status are categorical variables which were dummy coded. The error is represented by e.

Research Questions and Hypotheses

This study is guided by two research questions and six hypotheses:

> RQ1: Do the observed overall score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination predict the observed overall score on a mathematical state assessment?

$H_0$1: The observed overall score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do not predict the observed overall score on a mathematical state assessment.

H_a1: The observed overall score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do predict the observed overall score on a mathematical state assessment.

RQ2: Do the observed geometry domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination predict the observed score of the geometry domain on the state assessment?

$H_0$2: The observed geometry domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do not predict the observed geometry domain score on a mathematical state assessment.

H_a2: The observed geometry domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do predict the observed geometry domain score on a mathematical state assessment.

RQ3: Do the observed algebra domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic

status individually or in linear combination predict the observed algebra domain score on a mathematical state assessment?

$H_0$3: The observed algebra domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do not predict the observed algebra domain score on a mathematical state assessment.

H_a3: The observed algebra domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do predict the observed algebra domain score on a mathematical state assessment.

RQ4: Do the observed number sense domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination predict the observed score of the number sense domain on the state assessment?

$H_0$4: The observed number sense domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do not predict the observed number sense domain score on a mathematical state assessment.

H_{a4}: The observed number sense domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do predict the observed number sense domain score on a mathematical state assessment.

Threats to Validity

Five issues are covered in this section: internal and external validity, reliability, steps taken to mitigate these threats, and ethical procedures.

Internal Validity

Although there are many potential threats to research, this study is subject to three. Internal threats stem from procedures or treatments carried out by the researcher (Creswell, 2013). Threats may also come from the experiences of participants; experiences may skew the ability of the researcher to draw the correct inferences. Since the research data for this research comes from assessments, threats from participant experience are minimal. The collection of data from the sample population is a possible threat. Using archival data offers no control over the collection of data.

Another internal threat may come from omitted variable bias (Creswell, 2013). In the process of selecting variables for this study, numerous variables have been researched but some have been excluded because of time and the ability to include them. This threat may interfere with the interpretation of the results. In this research study, the demographic variables were selected in accordance with prior research; more predictive variables may be used in future research.

External Validity

External validity comes into question when researchers draw incorrect inferences from the data and apply them to a larger population (Creswell, 2013). I must guard against generalizing to other groups or populations, beyond what the research group representation allows. These threats arise because of participants' characteristics, such as

reactions to prior testing or bias experiences. Because this research uses archival data, there are few external threats. But one such threat is the generalizability of results in this study to people in other demographic or geographic settings.

Ethical Procedures

The two major considerations for the protection of individuals during this research are anonymity and confidentiality. These must be fulfilled for the safety of the participants. Although these might seem a minute harm, a violation of these is still a harm created by the researcher. Anonymity is created when de-identifying data is obtained from individuals (Frankfort-Nachmias & Nachmias, 2008). When a person cannot be specifically attached to the data, then the researcher has fulfilled this requirement. Walden's IRB approval was secured prior to obtaining and analyzing the archival data. The data I received is stored in my password protected personal home desktop that only I use. A back up file is stored in a personal external drive password protected folder that only I have access to. The data will be kept for five years after the study is concluded. In this research, I received data with students' names. After receiving all the students' data, I converted each student name into a unique number. The assigned number is the only manner to recall the student. Doing this ensures the anonymity of the individual.

Frankfort-Nachmias and Nachmias (2008) state that participants' information should remain confidential under most circumstances. In some cases, research confidentiality cannot be kept as an example of a subpoena from courts. In this research confidentiality is top priority. There is data such as state scores that are already public information. Detailed student information is not public information. Demographics and

geographic location of students whose data are used in this research are kept strictly confidential. Only group results have been made public.

Summary

The purpose of this research was to determine if there is a significant predictive relationship between district assessments and state assessments. I investigated the relationship as well as the strength of the relationship. With this information, a multiple regression model was used to predict student scores. As stated, the research questions in this study had two predictive variables, which are the district assessment scores, and the state assessment scores as the criterion variable. The demographic predictors for the multiple linear regression are the race, gender, and socioeconomic status of each student. The data was collected from assessments that have been taken by the students. The sample population is 250 students. The required minimum sample size was calculated to be 68 participants. However, to err on the side of caution, I randomly selected 100 participants, from the available population of 225 students. Multiple linear regression models were applied using their data. Students' names were converted into unique numbers in order to keep student's anonymity. Thus far, I have discussed and presented the research questions, hypotheses, literature review and methods of this study. The next step is to conduct the research study. In the next chapter, I review the data as well as report the results of the analysis.

Chapter 4: Results

Introduction

In this chapter, I discussed the process used to analyze the data and the results associated with each research question. The purpose of this research was to determine how well a student's midyear and end of the year assessment scores, race, gender, and socioeconomic status may predict the student's overall state assessment and domain scores. Four multiple linear regression models were implemented with the midyear and end of year assessment scores, race, gender, and socioeconomic status as predictors, and state domain score as the criterion. One model was used to predict the overall scores while the remaining three models were used to predict the respective scores in geometry, algebra, or number sense. The following are the research questions and their respective hypotheses for this research study:

RQ1: Do the observed overall score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination predict the observed overall score on a mathematical state assessment?

$H_0$1: The observed overall score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do not predict the observed overall score on a mathematical state assessment.

H_a1: The observed overall score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status

individually or in linear combination do predict the observed overall score on a mathematical state assessment.

RQ2: Do the observed geometry domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination predict the observed score of the geometry domain on the state assessment?

$H_0$2: The observed geometry domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do not predict the observed geometry domain score on a mathematical state assessment.

H_a2: The observed geometry domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do predict the observed geometry domain score on a mathematical state assessment.

RQ3: Do the observed algebra domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination predict the observed algebra domain score on a mathematical state assessment?

$H_0$3: The observed algebra domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do not predict the observed algebra domain score on a mathematical state assessment.

H_a3: The observed algebra domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do predict the observed algebra domain score on a mathematical state assessment.

RQ4: Do the observed number sense domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination predict the observed score of the number sense domain on the state assessment?

$H_0$4: The observed number sense domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do not predict the observed number sense domain score on a mathematical state assessment.

H_{a4}: The observed number sense domain score on a mathematical district midyear assessment, district end of year assessment, race, gender, and socioeconomic status individually or in linear combination do predict the observed number sense domain score on a mathematical state assessment.

Data Collection

The data for this study was received from archival records at a South Florida charter school. Written authorization was obtained from the principal of the school to use the existing data. The Walden IRB approval number is 09-22-20-0369279and the email from Walden IRB giving permission to begin this research is shown in Appendix B. After obtaining Walden IRB's approval for this study, I contacted the student records office to

request the data needed to conduct this research. The data was then transferred to me in Excel spreadsheets through a USB flash drive. To ensure anonymity, once I received the data, the names of students were replaced with unidentified numbers. I removed all cases in the dataset with missing data for any of the variables in this study. The assessment scores and demographics were then transferred into SPSS where a descriptive statistical analysis was conducted as well as testing for the four assumptions for multiple linear regression. Finally, four multiple linear regressions were conducted to examine the research questions and their respective hypotheses.

Sample

To determine the minimum sample size, a G*Power analysis was conducted. Using the software G*Power, the required minimum sample size was calculated to be 68 participants with a statistical power of .8 (Field, 2013). Although the minimum sample size was calculated to be 68 participants, to err on the side of caution, I randomly selected 100 participants from the available population of 225 students. Thus, the final sample size for this study was $n = 100$ students. In this sample 40% were White, 54% were male, and 56% did not have free or reduced lunch, as indicated in Table 1.

Table 1

Frequency of Student's Demographics

		Frequency	Percent	Valid Percent	Cumulative Percent
Race	White	40	40.0	40.0	40.0
	Non White	60	60.0	60.0	100.0
Gender	Male	54	54.0	54.0	54.0
	Female	46	46.0	46.0	100.0
Reduced Lunch	No	56	56.0	56.0	56.0
	Yes	44	44.0	44.0	100.0
	Total	100	100.0	100.0	

Descriptive Statistics of Sample

The average overall state score was 342.78 (*SD* = 17.757) while the average score for the state domain scores were: geometry 0.45 (*SD* = 0.194), algebra 0.467 (*SD* = 0.198), and number sense 0.509 (*SD* = 0.227), respectively. The average score for the midyear district assessment was 252.62 (*SD* = 9.811) while the average score for the midyear district domains were: geometry 225.00 (*SD* = 12.258), algebra 227.8 (*SD* = 10.623), and number sense 225.35 (*SD* = 15.016), respectively. The average score for the end of year district assessment was 225.1 (*SD* = 11.223) while the average score for the end of year district domains were: geometry 224.2 (SD = 12.765), algebra 226.4 (*SD* = 12.8), number sense 225.115 (*SD* = 16.875), respectively. More descriptive statistics are presented in Table 2.

Table 2

Descriptive Statistics of Assessments Score Variables

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
State Overall Score	100	99	294	393	342.78	17.757	315.305
State Geometry	100	.87	.07	.93	.4500	.19358	.037
State Algebraic Operations	100	.82	.12	.94	.4671	.19782	.039
State Number Sense	100	1.00	.00	1.00	.5090	.22746	.052
Midyear Overall Score	100	51	193	244	225.62	9.811	96.258
Midyear Geometry	100	70.0	185.5	255.5	225.000	12.2578	150.253
Midyear Algebraic Operations	100	60.0	195.5	255.5	227.800	10.6225	112.838
Midyear Number Sense	100	125.0	120.5	245.5	225.350	15.0161	225.482
End of Year Overall Score	100	62	187	249	225.10	11.223	125.949
End of year Geometry	100	60.0	195.5	255.5	224.200	12.7648	162.939
End of year Algebraic Operations	100	70.0	185.5	255.5	226.400	12.7995	163.828
End of year Number Sense	100	135.0	120.5	255.5	225.115	16.8752	284.772
Valid N (listwise)	100						

Assumptions of Multiple Linear Regression Models

Before preforming any of the multiple linear regression analyses, the assumptions needed for this analysis were tested. The assumptions of linearity, homoscedasticity, normality of distribution, and multicollinearity were assessed for each model in this research. The following are the explanations of each of the assumptions as well the results of their tests.

Linearity is established when the criterion variable has a linear relationship with the predictor variables (Field, 2013). The linear relationship was established by visual examination of each partial regression scatterplot for all multiple linear regression models. My examination indicated that all continuous predictor variables indicated a linear relationship. The partial regression scatterplots for the state overall assessment score model are displayed in Figure 1 and Figure 2, respectively.

Figure 1

Partial Regression Scatterplot of Midyear Overall Score by State Score

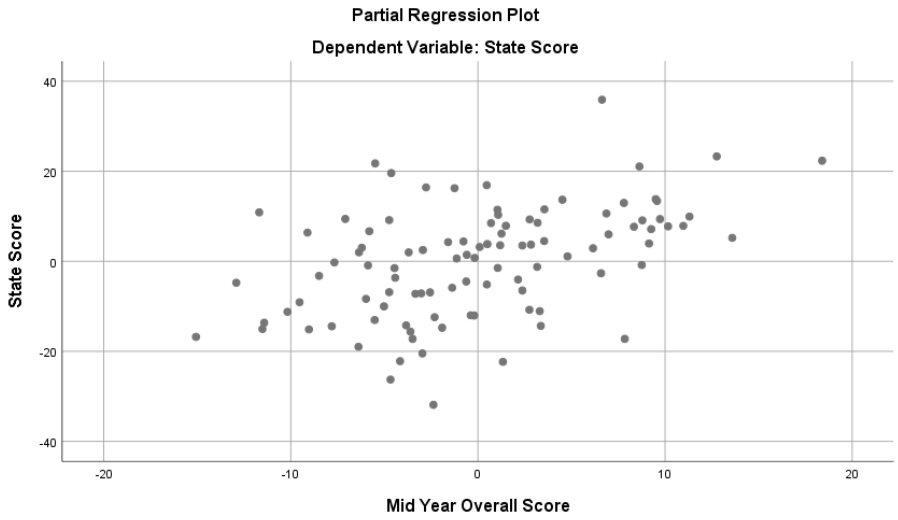
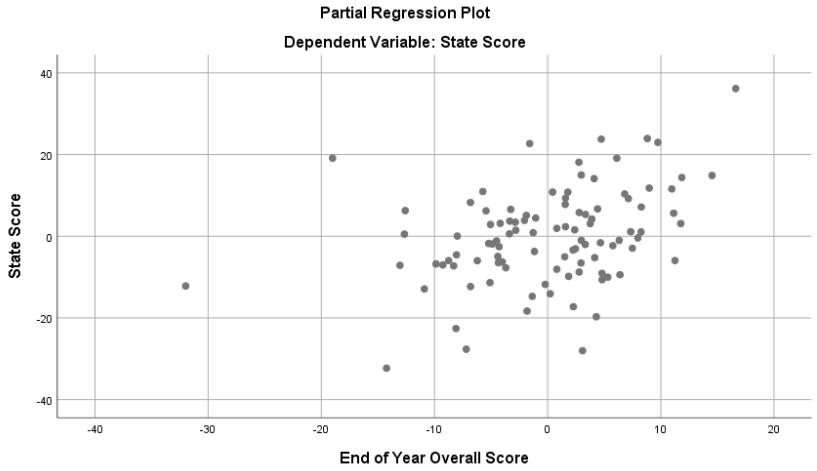


Figure 2

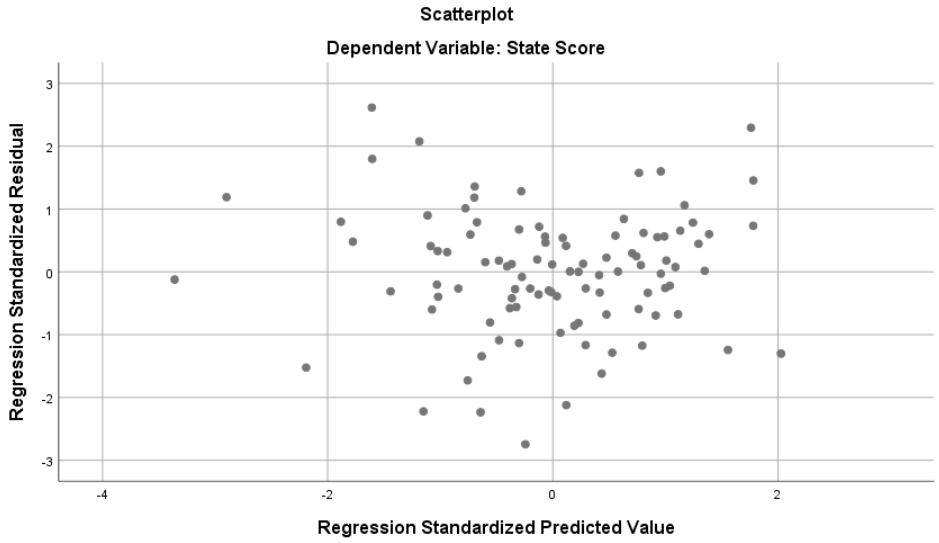
Partial Regression Scatterplot of End of Year Score by State



Homoscedasticity implies that the distribution across the predictor variables are homogeneous (Field, 2013). Scatterplots for each multiple linear regression model were used to examine homoscedasticity. The scatterplots indicated random distribution of data, which satisfies the assumption of homoscedasticity. All scatterplots also indicated all values being below 3 and above -3 which signifies that there were no outliers. Although there are four scatterplots, I illustrated the scatter plot that pertains to the main research question. The scatterplot for the state overall assessment score model is displayed in Figure 3.

Figure 3

Scatterplot of Regression Standardized Predicted Value by Regression Standardized Residual



Normality of distribution of scores was tested by visually examining the histogram and *p-p* plots for each multiple linear regression model. All four histograms indicated a robust correlation to the theoretical quartiles. All *p-p* plots indicated a strong tendency of the data towards a center line. Each model was found to be normally distributed. The histogram and *p-p* plot for Model 1 are displayed in Figure 4 and Figure 5, respectively.

Figure 4

Histogram of State Score Regression Standardized Residual by Frequency

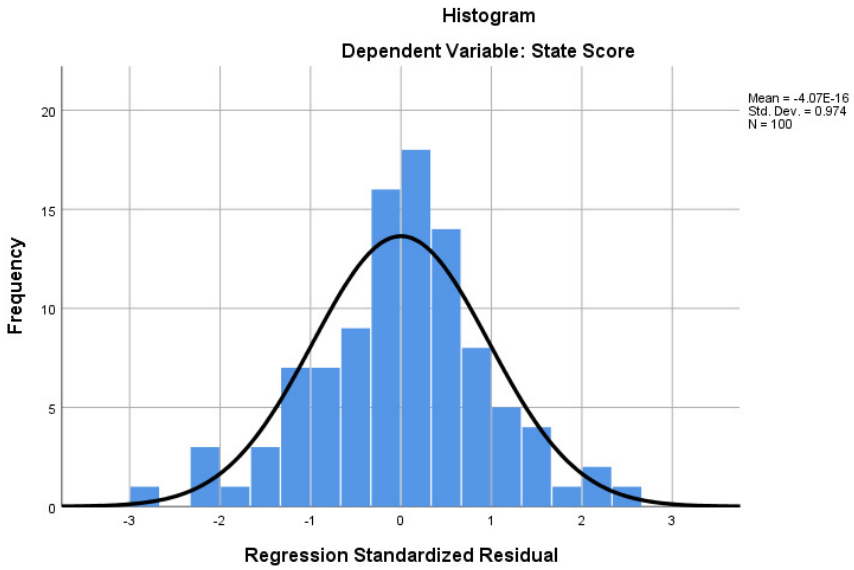
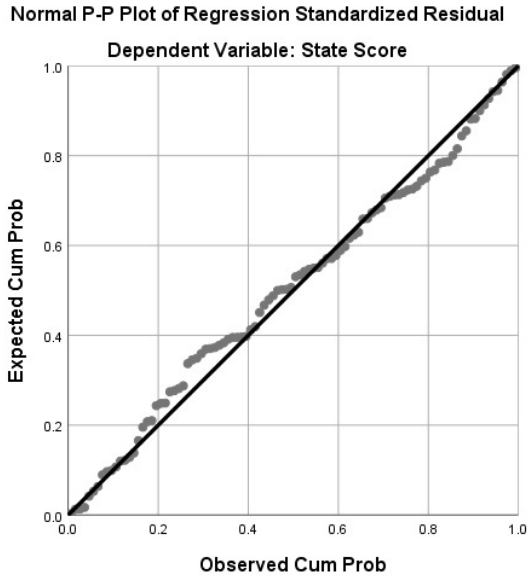


Figure 5

P-P Plot of State Score Regression Standardized Residual



Multicollinearity implies that there is a strong correlation between two or more independent variables (Field, 2013). VIF was used to assess multicollinearity. A VIF

value greater than 5 would be a cause for concern, while a VIF of 10 should be the maximum. (Field, 2013). All VIF values in this research were below 2.247, which indicates low concern. Although there are four models, I illustrated the table that pertains to the main research question. The collinearity data results for Model 1 are displayed in Table 3.

Table 3

Coefficients of State Score Criterion Variable

		Collinearity Statistics	
Model		Tolerance	VIF
1	Midyear Overall Score	.447	2.235
	End of Year Overall Score	.445	2.247
	Race	.925	1.081
	Gender	.959	1.043
	Reduced Lunch	.923	1.083

Data Analysis and Results

To answer each research question and its respective hypothesis, it was necessary to conduct one multiple linear regression model for each question. The first model analyzed the overall assessment scores, while the second, third, and fourth models analyzed the individual mathematical domains of geometry, algebra, and number sense. All demographic predictor variables (i.e., race, gender, and social economic status) were dummy coded. Race was coded White = 0 and Non-White was = 1. Similarly, gender was coded as male = 0 and female = 1. Finally, socioeconomic status of a student was coded as student paying full lunch = 0 and those on reduced or free lunch = 1. The data for all

four models are shown in Table 3. All models used a 95% confidence level to test each hypothesis.

Overall Scores

To examine Research Question 1, a multiple linear regression analysis was conducted to evaluate if the predictor variables of district midyear overall assessment scores, district end of year overall assessment scores, race, gender, and social economic status, individually or in linear combination, predicted the state assessment overall scores. Model 1 shows statistically significant results with F (5, 94) = 32.289, $p < .001$. The R^2 for model 1 was 0.632, indicating that more than 60% of the state assessment score was predicted by this model, while less than 40% comes from other factors. The complete results of all model summaries are displayed in Table 4.

Table 4

Model Summary

Model	R	R^2	Adjusted R^2	Std. Error of the Estimate	Change Statistics				
					R^2 Change	F Change	df1	df2	Sig. F Change
1[b]	.795[a]	.632	.612	11.054	.632	32.289	5	94	.000
2[d]	.564[c]	.318	.281	.16410	.318	8.753	5	94	.000
3[f]	.713[e]	.509	.483	.14227	.509	19.478	5	94	.000
4[h]	.503[a]	.254	.214	.20168	.254	6.384	5	94	.000

Note.
a. Predictors: (Constant), Midyear Score, End of Year Score, Race, Gender, Reduced Lunch
b. Dependent Variable: State Score
c. Predictors: (Constant), Midyear Geometry, End of year Geometry, Race, Gender, Reduced Lunch
d. Dependent Variable: State Geometry
e. Predictors: (Constant), Midyear Algebraic Operations, End of year Algebraic Operations, Race, Gender, Reduced Lunch
f. Dependent Variable: State Algebraic Operations
g. Predictors: (Constant), Midyear Number System, End of year Real Numbers, Race, Gender, Reduced Lunch
h. Dependent Variable: State Number Sense

The standardized *Beta* coefficients indicated the contribution of each predictor variable to the criterion of state assessment overall scores. The results of the first model revealed that race, gender, and socioeconomic status are not significant predictors of state assessment scores ($p > .05$). However, the results of the multiple linear regression analysis also revealed a significant relationship between district midyear assessment and the state assessment, as well as the district end of year and the state assessment. The midyear assessment score was found to be statistically significant with *Beta* = 0.866, 95% C.I. [.530,1.202], $p < .001$. Thus, for every unit of increase in midyear assessment score,

a 0.866 units increase in state assessment overall score was predicted, holding all other variables constant. The end of year assessment score was found to be statistically significant with *Beta* = 0.582, 95% C.I. [.287,.876], $p < .001$. Thus, for every unit of increase in end of year assessment score, a 0.582 units increase in state assessment overall score was predicted, holding all other variables constant. The complete results of all coefficient data are displayed in Table 5.

Table 5

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B	
		B	Std. Error	Beta	*t*	Sig.	Lower Bound	Upper Bound
1[a]	(Constant)	16.955	26.947		.629	.531	-36.548	70.458
	Midyear Overall Score	.866	.169	.478	5.115	.000	.530	1.202
	End of Year Overall Score	.582	.148	.368	3.920	.000	.287	.876
	Race	3.932	2.346	.109	1.676	.097	-.726	8.591
	Gender	.560	2.265	.016	.247	.805	-3.938	5.057
	Reduced Lunch	-2.159	2.318	-.061	-.932	.354	-6.761	2.443
2[b]	(Constant)	-1.728	.346		-4.995	.000	-2.415	-1.041
	Midyear Geometry	.004	.002	.229	2.271	.025	.000	.007
	End of year Geometry	.006	.002	.405	4.005	.000	.003	.009
	Race	.033	.035	.083	.934	.353	-.037	.102
	Gender	.028	.033	.073	.842	.402	-.038	.094
	Reduced Lunch	-.033	.035	-.085	-.954	.343	-.102	.036
3[c]	(Constant)	-2.687	.334		-8.034	.000	-3.351	-2.023
	Midyear Algebraic Operations	.008	.002	.433	4.997	.000	.005	.011
	End of year Algebraic Operations	.006	.001	.385	4.493	.000	.003	.009

	Race	.019	.030	.047	.621	.536	-.041	.079
	Gender	.022	.029	.056	.754	.453	-.036	.080
	Reduced Lunch	-.034	.030	-.086	-1.142	.257	-.093	.025
4[d]	(Constant)	-1.222	.381		-3.206	.002	-1.980	-.465
	Midyear Number Sense	.004	.001	.292	3.156	.002	.002	.007
	End of year Number Sense	.003	.001	.235	2.449	.016	.001	.006
	Race	.081	.043	.175	1.883	.063	-.004	.166
	Gender	-.092	.042	-.202	-2.179	.032	-.175	-.008
	Reduced Lunch	.009	.043	.020	.218	.828	-.075	.094

a. Dependent Variable: State Score
b. Dependent Variable: State Geometry
c. Dependent Variable: State Algebraic Operations
d. Dependent Variable: State Number Sense

This first multiple linear regression model was conducted to test the hypothesis of the first research question with the overall mathematical scores as the criterion variable. The results indicated a significant contribution to the prediction of the criterion variable by the midyear and end of year assessment scores predictor variables.

Geometry State Assessment Scores Analysis

To examine Research Question 2, a multiple linear regression analysis was also conducted to evaluate if the predictor variables of district midyear geometry domain scores, district end of year geometry domain scores, race, gender, and social economic status, individually or in linear combination, predict the state assessment geometry domain scores. Results of the analyses indicated the second model was statistically significant with $F(5, 94) = 8.753$, $p < .001$. The R^2 for model 2 was .318, indicating that

just over 31% of the state assessment geometry domain score was predicted by the model, while about 68% came from other factors.

The standardized *Beta* coefficients indicated that race, gender, and socioeconomic status are not significant predictors of state assessment scores ($p > .05$). However, the results of the multiple linear regression analysis revealed a statistically significant relationship between the midyear geometry domain score and the state geometry domain score. The regression coefficient *Beta* = 0.004, 95% C.I. [.000, .007], $p < .05$ associated with district midyear assessment geometry domain suggests that with each additional district midyear geometry domain assessment point earned, the geometry domain score on the state assessment tends to increase by approximately 0.004, while holding all other variables constant. Moreover, according to the model, the district end of year geometry assessment score was a significant predictor of the state geometry score, *Beta* = 0.006, 95% C.I. [.003, .009], $p < .001$. Therefore, for each additional end of year assessment point earned, the geometry domain score on the State assessment tends to increase by approximately 0.006, while holding all other variables constant. The results of this multiple linear regression model indicated a significant contribution to the prediction of the criterion variable by the midyear assessment geometry domain score and the end of year assessment geometry domain score predictor variables. Similar results were found in the following model.

Algebra State Assessment Scores Analysis

To examine research question 3, a multiple linear regression analysis was also conducted to evaluate if the predictor variables of district midyear algebra scores, district

end of year algebra scores, race, gender, and social economic status, individually or in linear combination, predict the state assessment overall scores. Model 3 indicates a statistical significance with $F(5, 94) = 19.478$, $p <. 001$. The R^2 for model 2 was .509, indicating that a little over 50% of the state assessment score was predicted by this model, while under 50% comes from other factors.

The standardized *Beta* coefficients indicated that race, gender, and socioeconomic status are not significant predictors of state assessment algebra domain scores ($p > .05$). The midyear assessment algebra domain score was found to be statistically significant with regression coefficient of *Beta* = 0.008, 95% C.I. [.050, .011], $p < .001$. Thus, for every unit of increase in midyear assessment algebra domain score, a 0.008 units increase in state assessment algebra domain score may be predicted, while holding all other variables constant. Additionally, the end of year assessment algebra domain score was found to be statistically significant with regression coefficient of *Beta* = 0.006, 95% C.I. [.003, .009], $p < .001$. Thus, for every unit of increase in end of year assessment algebra domain score, a 0.006 units increase in state assessment algebra domain score may be predicted, while holding all other variables constant. The results of this multiple linear regression model indicated a significant contribution to the prediction of the criterion variable by the midyear assessment algebra domain scores and the end of year assessment algebra domain scores predictor variables. Similar results were found for the research question 4.

Number Sense State Assessment Scores Analysis

To examine research question 4, a multiple linear regression analysis was also conducted to evaluate if the predictor variables of district midyear number sense domain scores, district end of year number sense domain scores, race, gender, and social economic status, individually or in linear combination, predict the state assessment number sense domain scores. Model 4 indicates a statistical significance with $F(5, 94) = 6.384$, $p < .001$. The R^2 for model 4 was .254, this indicates that just over 25% of the state assessment number sense domain score may be predicted by this model, while less than 75% comes from other factors.

The standardized *Beta* coefficients indicated that race, gender, and socioeconomic status are not significant predictors of state assessment number sense domain scores ($p > .05$). However, the results of the multiple linear regression analysis revealed a statistically significant relationship between the midyear number sense domain score and the state number sense domain score. The regression coefficient *Beta* = 0.004, 95% C.I. [.002, .007], $p < .05$ associated with district midyear assessment number sense domain suggests that with each additional district midyear number sense domain assessment point earned, the number sense domain score on the state assessment tends to increases by approximately 0.004, while holding all other variables constant. The regression coefficient *Beta* = 0.003, 95% C.I. [.001, .006], $p < .05$ associated with end of year assessment number sense domain suggests that with each additional end of year assessment point earned, the number sense domain score on the state assessment number sense domain tends to increase by approximately 0.003, while holding all other variables

constant. The results of this multiple linear regression model indicated a significant contribution to the prediction of the criterion variable by the midyear assessment number sense domain scores and the end of year assessment number sense domain scores predictor variables.

Summary

In this chapter, I presented the results of each research question and their corresponding hypotheses. The purpose of this study was to determine if there is a statistically significant relationship between district assessments and state assessments, the strength of this relationship and the predictability of the state assessments score. Multiple linear regression models, using the district midyear and end of year assessment scores as predictors and the state assessment scores as the criterion variable, were implemented. Out of the data received, a random sample of 100 participants was selected for the analysis in this research. The demographic information indicated 40% were White, 54% were male, and 56% did not have free or reduced. All multiple linear regression models in this research were tested and met the assumptions of linearity, homoscedasticity, normality, and multicollinearity.

The results of the analyses led to the rejection of all four null hypotheses. All four models indicated that the district midyear assessment scores and district end of year assessment scores do significantly predict the state overall and domain assessment scores. Thus, concluding that the data do not support any of the null hypothesis. Although the assessment predictor variables were found to have a significant relationship, the

demographic predictor variables (i.e., race, gender, and socioeconomic status) were shown to be not significant contributors in any of the state assessment scores.

In the following chapter, I will present the interpretation and explanation of the data results given in this chapter. The implications to the literature review will also be discussed. Finally, recommendations and final conclusions of this study will also be given.

Chapter 5 Discussion, Conclusion, and Recommendations

Introduction

The purpose of this quantitative study was to determine if there is a statistically significant relationship and strength of the relationship between district assessments scores and state assessments scores. The assessments examined in this research are based on the current national standards. The assessment questions are constructed to measure the level of student mastery of a standard (CCSS Initiative, n.d.). Main (2012) reviewed the timing of introduction of new material, the lack of established curriculum, and the need for professional development to prepare teachers. Main advised that the CCSS should be subjected to examination, trials, and revisions by educational practitioners. Due to the gap created by changing educational policies as the common core, it is important we consistently review student scores and what variables significantly affect achievement.

The results from the multiple linear regression analyses indicated that the district midyear and the end of year assessment overall scores do significantly predict the state assessment scores. Respectively the geometry, algebra, and number sense domain scores from the district midyear and end of year assessment significantly predict the state assessment geometry, algebra, and number sense domain scores. However, the demographic predictor variables of race, gender, and socioeconomic status do not significantly contribute in predicting any of the state assessment scores. In this chapter, the interpretations of the findings, limitations, recommendations, and implications will be discussed.

Interpretation of the Findings

There were three major findings in this study that are of interest. The first finding pertains to the first research question which included the predictability of the district assessments overall scores to the state assessment overall scores. For this research question, the data indicated statistically significant results with F (5, 94) = 32.289, $p <$ .001, and $R^2 = 0.632$, Thus indicating that more than 60% of the state assessment score was predicted by the variables in the model. The midyear assessment score was found to be statistically significant with *Beta* = 0.866, 95% C.I. [.530,1.202], $p <$ .001. This indicates that for every unit of increase in midyear assessment score, a 0.866 unit increase in state assessment overall score can be predicted. Similarly, the end of year assessment score was found to be statistically significant with *Beta* = 0.582, 95% C.I. [.287,.876], $p <$ .001. Thus, for every unit of increase in end of year assessment score, a 0.582 unit increase in state assessment overall score may be predicted. The results indicate that out of all the predictive variables in this study the midyear overall score has the most impact on the state overall score.

The second finding of interest pertains to the mathematical domains of geometry, algebra, and number sense. These findings are the results of RQ 2, 3, and 4, respectively. Groß et al. (2016) and Shivraj (2017) both indicated the importance of additional research on mathematical domains. Shivraj noted that these domains are tested in several international assessments. Since the beginning of the CCSS implementation, the data pertaining to these domains are of high importance.

For geometry, the model indicated statistically significant results with $F(5, 94) = 8.753$, $p < .001$, and R^2 was .318, indicating that just over 31% of the state assessment geometry domain score can be predicted by the variables in the model. The regression coefficient for the midyear assessment was *Beta* = 0.004, 95% C.I. [.000, .007], $p < .05$. Similarly, the regression coefficient for the end of year assessment was *Beta* = 0.006, 95% C.I. [.003, .009], $p < .001$. Thus, both assessments are significant contributors to the state assessment scores.

Results of the analysis to predict the algebra state assessment score indicated a statistically significant model with $F(5, 94) = 19.478$, $p <. 001$, and $R^2 = .509$, which indicated that about 50% of the state assessment score may be predicted by this model. The midyear assessment algebra domain score was found to be statistically significant with regression coefficient of *Beta* = 0.008, 95% C.I. [.050, .011], $p < .001$. Moreover, the end of year assessment algebra domain score was also found to be statistically significant with regression coefficient of *Beta* = 0.006, 95% C.I. [.003, .009], $p < .001$. Just as with the geometry domain, both the midyear and end of year assessments are significant contributors in predicting the state assessment scores.

Similar findings were obtained regarding the number sense domain. This model indicated a statistical significance with $F(5, 94) = 6.384$, $p < .001$, and $R^2 = .254$. Thus, indicating that just over 25% of the state assessment number sense domain score may be predicted by this model. The midyear assessment number sense domain score was found to be statistically significant with regression coefficient of *Beta* = 0.004, 95% C.I. [.002,

.007], $p < .05$. The end of year number sense domain score regression coefficient was *Beta* = 0.003, 95% C.I. [.001, .006], $p < .05$.

The results in this research indicated that the district midyear domain scores and district end of year domain scores both were significant contributors to all three state assessment domain scores in this research. However, the domain with the strongest relationship to the state assessment score is algebra. This appears to align with previous data. Cipriani (2015) presented a detailed history of what are now the national math standards. In 2009, the National Governors Association and the Council of Chief State School Officers set the goal to create a framework to prepare students for college and work as well as have consistent standards nationwide. Of all the mathematical domains, algebra was determined to be the most substantial domain. The standards in the algebra domain were created to be vertical for a long-term vision (Lee, 2016). Since then, algebra has taken a much bigger part in math classes. The data in this current research, which indicated the algebra domain as the strongest predictive relationship, aligns with the results of these prior research on algebra.

In another study conducted by Jian-Hua Liang et al. (2018), the predictive power of cognitive and noncognitive variables to predict performance of eight graders on algebra state assessment was researched. Like the results of my research, their results indicated that the algebra scores had the most significance to the eighth-grade state scores, it accounted for 61% of explained variance.

The third finding of interest in this research is the findings on the demographic predictor variables of race, gender and socioeconomic status and their nonsignificance in

predicting the state assessment scores. Gottfried (2016) conducted research like my research on which Gottfried investigated the role of gender, race, and socioeconomic status on student end of year mathematical achievement. That study indicated that both males and females had gains from one assessment to another. It was also indicated that Whites and Blacks had a significant difference, while Hispanics and Asians did not have a significant difference. In that same study, socioeconomic status was significant for the students who had a lower socioeconomic status. The study conducted by Jian-Hua Liang et al. (2018) indicated that there was a difference in race and achievement between White, Asian, Hispanic/Latino, African American, and American Indian. The findings of my research indicated that the demographic predictor variables of race, gender, and socioeconomic status have no predictive significance in any of the state assessment scores. This means that in this population and geographic location these demographics are not significant predictive factors of mathematical state assessment score.

Limitations of the Study

In this study only eighth grade students were included. The reason for this inclusion was to follow the inclusion of this grade in research by NCES which is part of the U.S. Department of Education. The NAEP math assessments are given to fourth, eighth, and 12th grades (NCES, n.d.). The selection of eighth grade data is to have the option to align this research with the NAEP scores later. For the reason of time and resources, it was necessary to exclude other grades. Although this was a conscious and thought-out selection, it is a delimitation of this study.

As stated in in Chapter 1, the biggest limitation of this research is data itself. The assessment was administered by the local school. The data was then received from the school's students' records office. I did not create or collect the data myself; this creates high dependency on assessment administrators, students, and school's student records office to act in good faith.

Recommendations

After reviewing prior research and the data from this current research, several recommendations are suitable. The first recommendation would be to repeat this current research in a different geographic location. This research was conducted in a suburban charter school; it would be beneficial to see the results from different districts and populations across the United States. I would also recommend public and private schools that may administer state assessments be included to know whether the same results are reached. Conducting this research in urban and rural populations would also be beneficial to the body of knowledge.

The second recommendation would be to research more academic years. One of the limitations of this current research is that it was only conducted with an eighth-grade population. Although this was done in the effort to preserve time and resources, there are many other grades that should be studied. The NAEP math assessments are given to fourth, eighth, and 12th grades. It would be beneficial to include fourth and 12th grade student population in this study. This may lead to an opportunity to compare the results of this study to the research conducted by the NAEP.

The third recommendation would be to repeat this study and include more student demographics. In this study the demographic predictor variables included were race, gender, and socioeconomic status. VanDerHeyden et al. (2017) recommend to always include demographics of student in assessment research. This recommendation is the standard for similar studies. What does vary is what is included as demographic variables. The variables in this study are essential but there are additional demographics that may be included such as English as a second language, parental education level, parental annual income, number of siblings, parental involvement, and student living situation. The reason for the exclusion of these demographics from the current study is that some of these demographics may be difficult to attain but these would substantially add to the body of knowledge. Additionally, in this current study the categories for race were White and non-White. This was done because of the small number of participants other than White. I would recommend increasing the sample to be able to analyze specific races other than White and non-White.

Lastly, the fourth recommendation would be to repeat this study with other subjects. This study was conducted for the mathematics state assessment; however, the national standards were written for K-12 English and mathematics courses and supplement 6-12 history, social studies, science, and technical subjects (CCSS Initiative, n.d.). As I research studies like this current research, such as Jian-Hua Liang et al. (2018), in addition to mathematics their research included English assessment results. The template of this current study can be administered to any subjects with district and state

assessments. This recommendation would be very fitting especially to the English assessments that are just as high stakes as the mathematics assessments.

Although the results of my research were accurate and representative of the population chosen, the recommendation for additional studies would be fitting. As with prior research recommendations, it is the recommendation of this research to continually add research with current results.

Implications

Based on current educational policies, the findings from this study have much potential impact and implications for social change. The purpose of this study was to determine if there was a significant relationship between the district assessments, demographic variables, and the state assessment. The results from this research may potentially positively impact school districts, school administration, teachers, students, and parents. Some of these stakeholders may benefit more directly than others. As stated in Chapter 1, the results of this study give current data for lawmakers as well as state and district administration to make decisions regarding assessments and student achievement of standards. The results of this research will also give school districts credibility in continuing the administration of these assessments. Main (2012) recommended more data collection and research to provide a better picture of the assessments and current standards. Although this research did not assess how the current standards are being implemented, the results demonstrate the predictability of student achievement on state standards from assessments throughout the year. The results of this data give credibility to the administration of district assessments. Teachers may be impacted directly as they

may follow these relationships to increase student state scores. The biggest benefactors of this research are students since it is their scores that can be positively increased. Lastly, this information may be valuable to parents in knowing the variables that help increase scores on their child's state assessment.

In reviewing prior research, there are studies that indicate varying results. There has been a divided and critical opinion of the common core educational policy (Main, 2012). The opposition to the common core has come from the lack of information and explanation on how the standards would improve education (McGuinn, 2015). My current research may help shed light on the common core debate with up to current data. The findings may also provide examples of what is working. Other schools may emulate what is working within the chosen population. Policy makers are not necessarily educational experts but may use the data in this research to assist in making policy decisions. The impact of this educational research may reach beyond the common core standards into future policies. Lastly, the demographics included in this study were not significant predictors of state assessments. This implies that race, gender, or socioeconomic status of a student would not affect the opportunities to increase their scores. The results of this data would be very valuable for our society to be informed of.

Conclusion

This study was conducted to find out if there was a statistically significant relationship and strength of the relationship between district assessments scores as well as the demographics of race, gender, and socioeconomic status to state assessments scores. Although there has been prior research on predictors of mathematical assessment scores

as well as demographic impact, there have been varying results. For this reason, those same studies have called for more research. The goal of this research was to fill the gaps in data as well as to add to the existing literature regarding assessments that fall under the CCSS.

The findings of this study indicated that district assessments have a significant relationship to state assessments. The results also indicated that the demographics included in this study had no significance in scores. As education moves forward, these results may provide needed support to invested stakeholders.

References

Akkus, M. (2016). The Common Core State Standards for Mathematics. *International Journal of Research in Education and Science, 2*(1), 49–54.

Andersson, C., & Palm, T. (2017). The impact of formative assessment on student achievement: A study of the effects of changes to classroom practice after a comprehensive professional development programme. *Learning and Instruction, 49,* 92–102. https://doi-org.ezp.waldenulibrary.org/10.1016/j.learninstruc.2016.12.006

Anselmo, G. A., Yarbrough, J. L., Kovaleski, J. F., & Tran, V. V. N. (2017). Criterion-related validity of two curriculum-based measures of mathematical skill in relation to reading comprehension in secondary students. *Psychology in the Schools 54*(9), 1148–1159. https://doi-org.ezp.waldenulibrary.org/10.1002/pits.22050

Ateh, C. M., & Wyngowski, A. J. (2015). The Common Core State Standards: An opportunity to enhance formative assessment in history/social studies classrooms. *Clearing House, 88*(3), 85–90. https://doi-org.ezp.waldenulibrary.org/10.1080/00098655.2015.1023245

Banks, A., LaFors, J., & Education Trust-West. (2015). *Changing the equation: Ensuring the common core math standards enable all Students to excel in California schools. K-12 practice.* . Education Trust-West

Bitter, J. M. (2018). Creating a constructivist online music course for community college students: Theory and practice. *Community College Enterprise, 24*(2), 55–67.

Bohrnstedt, G., Kitmitto, S., Ogut, B., Sherman, D., Chan, D., National Center for Education Statistics (ED), & American Institutes for Research. (2015). *School composition and the Black-White achievement gap. NCES 2015-018. National Center for Education Statistics.* National Center for Education Statistics.

Bozkurt, G. (2017). Social constructivism: Does it succeed in reconciling individual cognition with social teaching andlearning practices in mathematics? *Journal of Education and Practice, 8*(3), 210–218.

Charter Schools USA (n.d.). *Home*. https://www.charterschoolsusa.com/

Choi, K. M., Lee, Y.-S., & Park, Y. S. (2015). What CDM can tell about what students have learned: An analysis of TIMSS eighth grade mathematics. *EURASIA Journal of Mathematics, Science & Technology Education, 11*(6), 1563–1577.

Cipriani, P. (2015). The evolution of mathematics standards: Have Common Core State Standards ended the math wars? *Mid-Atlantic Education Review, 3*(1), 2–8.

Common Core State Standard Initiative (n.d.). *About the standards*. http://www.corestandards.org/about-the-standards/

Cpalms. (n.d.) *Standards*. *http://www.cpalms.org/Public/search/Standard#0*

Creswell, J. W. (2013). *Research Design: Qualitative, quantitative, and mixed methods approaches* (Laureate Education, Inc., custom ed.). Sage Publications, Inc.

Daro, P., & Burkhardt, H. (2012). A population of assessment tasks. *Journal of Mathematics Education at Teachers College, 3*(1), 19–25.

Deane, P., Sabatini, J., Feng, G., Sparks, J., Song, Y., Fowles, M., & Foley, C. (2015). Key practices in the English Language Arts (ELA): Linking learning theory,

assessment, and instruction. Research Report. ETS RR-15-17. *ETS Research Report Series.*

Dewey, J. (1910). Concrete and Abstract Thinking. In *How we think* (pp. 135-144). D C Heath. https://doi-org.ezp.waldenulibrary.org/10.1037/10903-010

Exstrom, M., & Thatcher, D. (2014). Common core: Put to the test. *State Legislatures, 40*(8), 22-25.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Sage Publications.

Flores-Koulish, S. A., & Smith-D'Arezzo, W. M. (2016). "The three pigs": Can they blow us into critical media literacy old school style? *Journal of Research in Childhood Education, 30*(3), 349–360.

Frankfort-Nachmias, C. & Nachmias, D. (2008). *Research methods in the social sciences, (7th ed.).* Worth Publishers.

Gershenson, S., & Hayes, M. S. (2017). The summer learning of exceptional students. *American Journal of Education, 123*(3), 447–473. https://doi-org.ezp.waldenulibrary.org/10.1086/691226

Gibson, J. W. (2013). *The relationship between critical thinking, parent education, family income, extracurricular activity and intrinsic motivation in public schools: A multiple regression analysis* (Order No. 3598462) [Doctoral dissertation, Capella University]. ProQuest Dissertations & Theses Global

Gottfried, M. (2016). The role of real-life mathematics instruction on mathematics outcomes in kindergarten. *British Educational Research Journal, 42*(2), 314–341. https://doi-org.ezp.waldenulibrary.org/10.1002/berj.3207

Grady, M., Watkins, S., & Montalvo, G. (2012). The effect of constructivist mathematics on achievement in rural schools. *Rural Educator*, *33*(3), 37-46.

Green-Davis, Sha-Rhonda, M. (2017). *Multiple Roles as Predictors of Subjective Well-Being in African American Women.* ScholarWorks.

Groß, J., Robitzsch, A., & George, A. C. (2016). Cognitive diagnosis models for baseline testing of educational standards in math. *Journal of Applied Statistics, 43*(1), 229.

Harkness, L. M. (2016). *The effect of a constructivist-based approach on fifth grade reading achievement* (Order No. 10112034) [Doctoral Dissertation, Walden University]. ProQuest Dissertations & Theses Global

Harman, W. G., Boden, C., Karpenski, J., & Muchowicz, N. (2016). No Child Left Behind: A postmortem for Illinois. *Education Policy Analysis Archives, 24*(48).

Henderson, S., Petrosino, A., Guckenburg, S., Hamilton, S., & Northeast and Islands Regional Educational Lab. at Brown Univ., P. R. (2007). *Measuring How Benchmark Assessments Affect Student Achievement. Issues & Answers. REL 2007-No. 039. Regional Educational Laboratory Northeast & Islands.* Regional Educational Laboratory Northeast & Islands.

Hornbeck, D. (2017). Seeking civic virtue: *Two Views of the Philosophy and History of Federalism in U.S. Education. Journal of Thought, 51*(3–4), 52.

Houseworth, L. M. (2015). *Extended instruction in business courses to enhance student achievement in math* (Order No. 3687689) [Doctoral Dissertation, Walden University]. ProQuest Dissertations & Theses Global.

James, M. (2006). Assessment, teaching and theories of learning. In J. Gardner (Ed.), *Assessment and learning* (pp. 47-60) Sage https://doi.org/10.13140/2.1.5090.8960

Jian-Hua Liang, Heckman, P. E., & Abedi, J. (2018). Prior year's predictors of eighth-grade algebra achievement. *Journal of Advanced Academics, 29*(3), 249–269. https://doi-org.ezp.waldenulibrary.org/10.1177/1932202X18770172

Kretchmar, J. (2015). Progressive education. *Research Starters: Education (Online Edition),*

Lee, J. (2016). Implementing college and career standards in math methods course for early childhood and elementary education teacher candidates. *International Journal of Science and Mathematics Education, 14*(1), 177–192.

Louisiana State Dept. of Education, B. R. (2013). *Common Core State Standards for Mathematics: Teacher Self-Learning Series. Module 2: Focus and Coherence--The First Two CCSSM Shifts.* Louisiana Department of Education.

Main, L. (2012). Too much too soon? Common core math standards in the early years. *Early Childhood Education Journal, 40,* 73-77. doi:10.1007/s10643-011-0484-7.

Marinho, P., Leite, C., & Fernandes, P. (2017). Mathematics summative assessment practices in schools at opposite ends of performance rankings in Portugal. *Research in Mathematics Education, 19*(2), 184.

Markowitz, A. J. (2018). Changes in School Engagement as a Function of No Child Left Behind: A comparative interrupted time series analysis. *American Educational Research Journal, 55*(4), 721–760.

McCracken, B. (2014). Common core from the outside looking in. *Knowledge Quest, 42*(3), 8-9.

McDuffie, A. R., Drake, C., Choppin, J., Davis, J. D., Magana, M. V., & Carson, C. (2015). Middle School mathematics teachers' of the common core state standards for mathematics and related assessment and teacher evaluation Systems. *Educational Policy, 31*(2), 139–179. https://doi-org.ezp.waldenulibrary.org/10.1177/0895904815586850

McGuinn, P. (2015). Complicated politics to the Core: the bumpy implementation of the Common Core and its assessments has a thorny knot of political alliances and adversaries. *Phi Delta Kappan, 97*(1), 14.

Nahornick, A. (2016). Math Matters: Comparison of the mathematics requirements for bachelor of arts and bachelor of science degrees in Canada. *Canadian Journal of Higher Education,46*(1), 109-120.

National Center for Education Statistics. (n.d.). Frequently asked questions. https://nces.ed.gov/nationsreportcard/about/faqs.aspx

Nichols, J. B. (2017). Piaget and the common core: Is it the "American question?" *International Journal of Pedagogy & Curriculum, 24*(1), 17–25.

Reel, L. A. (2010). *The impact of unlimited access to graphing calculators on mathematics achievement of low socioeconomic students* (Order No. 3403827)

[Doctoral Dissertation, Walden University]. ProQuest Dissertations & Theses Global.

Rentner, D. S., Ferguson, M., & Center on Education, P. (2014). A research agenda for the common core state standards: What information do policymakers need? *Center On Education Policy, 4*

Richardson, C. J., & Eddy, C. M. (2011). The mathematical argument. *American Educational History Journal, 38*(1/2), 277.

Robbins, J. (2013). Uncommonly bad. *Academic Questions, 26*(1), 8-19. doi:10.1007/s12129-013-9336-9.

Schoenfeld, A. H. (2015). Summative and formative assessments in mathematics supporting the goals of the common core standards. *Theory Into Practice 54*(3), 183–194. https://doi-org.ezp.waldenulibrary.org/10.1080/00405841.2015.1044346

Sforza, D., Tienken, C. H., & Kim, E. (2016). A comparison of higher-order thinking between the common core state standards and the 2009 New Jersey content standards in High School. *AASA Journal of Scholarship & Practice, 12*(4), 4.

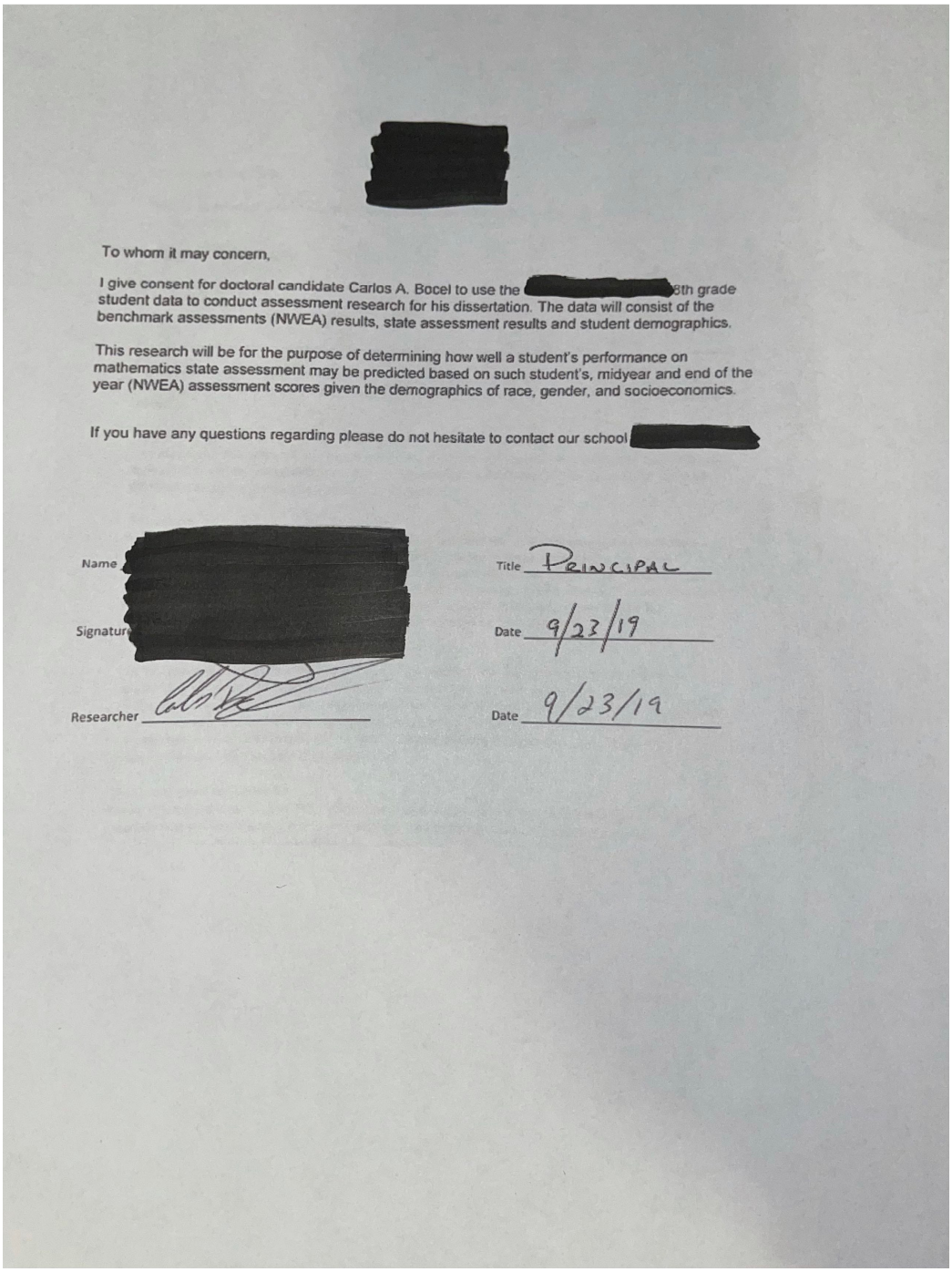
Shivraj, P. (2017). Evaluating the (mis)alignment of the intended to the assessed curriculum for the U.S.: Implications for the common core state standards for mathematics. *International Journal of Education in Mathematics, Science and Technology, 5*(4), 333–347.

Tampio, N. (2017). Democracy and national education standards. *Journal of Politics, 79*(1), 33–44. Retrieved from https://doi-org.ezp.waldenulibrary.org/10.1086/687206

Traynor, A. (2017). Sensitivity of cross-state assessment item difficulty to differences in state curricular content standards. *Applied Measurement in Education, 30*(4), 329–342. https://doi-org.ezp.waldenulibrary.org/10.1080/08957347.2017.1353991

VanDerHeyden, A. M., Codding, R. S., & Martin, R. (2017). Relative value of common screening measures in mathematics. *School Psychology Review, 46*(1), 65–87.

Appendix A: Data use Agreement



[redacted]

To whom it may concern,

I give consent for doctoral candidate Carlos A. Bocel to use the [redacted] 8th grade student data to conduct assessment research for his dissertation. The data will consist of the benchmark assessments (NWEA) results, state assessment results and student demographics.

This research will be for the purpose of determining how well a student's performance on mathematics state assessment may be predicted based on such student's, midyear and end of the year (NWEA) assessment scores given the demographics of race, gender, and socioeconomics.

If you have any questions regarding please do not hesitate to contact our school [redacted]

Name [redacted] Title PRINCIPAL

Signature [redacted] Date 9/23/19

Researcher _____ Date 9/23/19

Appendix B: IRB's Permission and Number

IRB Materials Approved - Carlos Bocel

[redacted]

This email is to notify you that the Institutional Review Board (IRB) confirms that your study entitled, "District Scores and Demographics as Predictors of State Mathematics Assessment Scores," meets Walden University's ethical standards. Our records indicate that you will be analyzing data provided to you by Coral Springs Charter School as collected under its oversight. Since this study will serve as a Walden doctoral capstone, the Walden IRB will oversee your capstone data analysis and results reporting. The IRB approval number for this study is 09-22-20-0369279, which expires when your student status ends.

This confirmation is contingent upon your adherence to the exact procedures described in the final version of the documents that have been submitted t[redacted]f this date. This includes maintaining your current status with the university and the oversight relationship is only valid while you are an actively enrolled student at Walden University. If you need to take a leave of absence or are otherwise unable to remain actively enrolled, this is suspended.

If you need to make any changes to your research staff or procedures, you must obtain IRB approval by submitting the IRB Request for Change in Procedures Form. You will receive confirmation with a status update of the request within 1 week of submitting the change request form and are not permitted to implement changes prior to receiving approval. Please note that Walden University does not accept responsibility or liability for research activities conducted without the IRB's approval, and the University will not accept or grant credit for student work that fails to comply with the policies and procedures related to ethical standards in research.

When you submitted your IRB materials, you made a commitment to communicate both discrete adverse events and general problems to the IRB within 1 week of their occurrence/realization. Failure to do so may result in invalidation of data, loss of academic credit, and/or loss of legal protections otherwise available to the researcher.

Both the Adverse Event Reporting form and Request for Change in Procedures form can be obtained at the Documents & FAQs section of the Walden web site: [redacted]

Doctoral researchers are required to fulfill all of the Student Handbook's Doctoral Student Responsibilities Regarding Research Data regarding raw data retention and dataset confidentiality, as well as logging of all recruitment, data collection, and data management steps. If, in the future, you require copies of the originally submitted IRB materials, you may request them from Institutional Review Board.

Both students and faculty are invited to provide feedback on this IRB experience at the link below:

[redacted]

Information about the Walden University Institutional Review Board, including instructions for application, may be found at this link [redacted]